The Dissertation Committee for Gregory James Hatlestad Certifies that this is the approved version of the following dissertation:

Characterization of Glabra2 and Transparent Testa Glabra2, Targets of the TTG1 Complex

Committee:

Alan Lloyd, Supervisor

Enamul Huq

Thomas Juenger

Mona Mehdy

Stanley Roux

Characterization of Glabra2 and Transparent Testa Glabra2, Targets of the TTG1 Complex

by

Gregory James Hatlestad, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August, 2010

Dedication

To My Family.

Acknowledgements

No achievement can be accomplished without the aid of others. I would like to acknowledge the aid and support supplied by my committee members and the past and present members of the Lloyd lab especially Tony Gonzalez. I would also like to thank Alan Lloyd for the support and guidance to be able to accomplish this work.

Characterization of Glabra2 and Transparent Testa Glabra2, Targets of the TTG1 Complex

Publication No._____

Gregory James Hatlestad

The University of Texas at Austin, 2010

Supervisor: Alan Lloyd

Studies on epidermal cell fate determination have been important for gaining insight into the genetic and molecular mechanisms leading to the differentiation and patterning of cells. In *Arabidopsis,* the organization and development of many epidermal characters including trichomes, root hairs and the seed coat have been found to be controlled by a single combinatorial transcription factor complex consisting of a WD-repeat containing protein, Transparent Testa Glabra 1 (TTG1), and various MYB and bHLH proteins. The work here consists of identification of Glabra2 (GL2) and Transparent Testa Glabra2 (TTG2) as direct transcriptional targets of the TTG1 combinatorial complex, further characterization of GL2 function, and identification of transcriptional targets of GL2 and TTG2. Both GL2 and TTG2 are important in the regulation of trichomes, root hairs and seed coat development.

GL2 has been identified as an important regulator of epidermal cell fate for over fifteen years yet there is little known about its function and only three transcriptional targets are identified, all involved in root hair patterning. Through the examination of its function a nuclear localization signal was verified and shown that GL2 homodimerizes. Through analysis of available expression databases and differential sequence analysis using SOLiD sequencing technology, several direct targets of GL2 and many more possible transcriptional targets of both GL2 and TTG2 were identified in trichomes. Some of these targets are members of the TTG1 complex, and they are all specialized in the maturation of trichomes, suggesting that GL2 switches the focus of the complex by activating the TTG1 complex members involved in maturation of the trichome through a feedback mechanism.

Examination of *gl2* mutants shows that they do not produce trichome accessory cells which usually surround the trichome. An additional target of GL2 is At5g65300 which when overexpressed results in the elongation and proliferation of trichome accessory cells into a tall pillar of cells. This suggests that GL2 is involved in the regulation of accessory cell development through At5g65300.

The work presented here represents important advances of our knowledge of epidermal cell fate through characterization of the major downstream regulators of epidermal development.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Studies on epidermal cell fate determination have been important for gaining insight into the genetic and molecular mechanisms leading to the differentiation and patterning of cells. In *Arabidopsis,* the organization and development of many epidermal characters including trichomes, root hairs and the seed coat have been found to be controlled by a single combinatorial transcription factor complex consisting of a WD-repeat containing protein, Transparent Testa Glabra 1 (TTG1), and various MYB and bHLH proteins (Oppenheimer et al. 1991; Walker et al. 1999; Payne et al. 2000; Zhang et al. 2003). This complex has also been implicated in the development of cotton fibers thus granting possible commercial benefits to continued research into this transcriptional complex and the pathways it regulates (Lee et al. 2007). Intensive investigation has generated a large amount of information about TTG1 dependent pathways. This includes the fact that certain members of the complex itself are transcriptional targets of the complex, i.e. the complex is self regulating. The complex's main targets in the regulation of development pathways are a homeodomain-START locus, Glabra2 (GL2) and a WRKY locus, Transparent Testa Glabra2 (TTG2) (Morohashi et al. 2007; Zhao et al. 2008). Mutations in GL2 result in changes in trichome, root hair and seed coat development that resemble mutations in the TTG1 complex members (Rerie et al. 1994; Szymanski et al. 1998) while mutations in TTG2 have defects in trichomes and seed coat development. However, except for two repressed and one activated target in the root hair patterning pathway, virtually nothing is known about how GL2 regulates development (Ohashi et al. 2003; Tominaga-Wada et al. 2009). There is currently nothing known

about how TTG2 regulates development. In order to further our understanding of plant epidermal development, I have performed numerous experiments aimed towards understanding the functions of GL2 and TTG2 with particular emphasis on the identification of transcription targets in the trichome cell fate and differentiation pathway.

The TTG1 dependent pathways include the cell fate and differentiation of trichomes, root hairs, and outer seed coat cells, while also including the flavonoid biosynthetic pathways such as the production of anthocyanins and tannins. In Arabidopsis, trichomes are single celled projections on the leaf surface with several branch points. They are evenly distributed across the leaf rarely occurring adjacently. Root hairs are also single cell projections on the root yet they are patterned much more strictly. Root hairs occur in vertical files with the root hair file forming in epidermal cells which are in contact with two cortex cells. In seed development, the TTG1 complex, GL2, and TTG2 regulate the differentiation of the outer layer of the seed coat, particularly the production and disposition of mucilage resulting in the formation of the columella in those cells. The TTG1 complex and TTG2 also control the production of seed coat tannins in the inner testa layer.

TTG1 is a WD repeat containing protein that along with various MYBs and bHLHs, forms a combinatorial complex. While the MYBs and bHLHs change depending on which cell fate is in question, TTG1 is required in all cases (Fig 1.1). There have been 9 MYB proteins and 4 bHLH proteins implicated in the TTG1 dependent pathways. The MYBs are very similar, yet for each pathway there is a different MYB that serves as the major complex member. Glabra1 (GL1) is the major MYB (Oppenheimer et al. 1991) for trichomes, Werewolf (WER) is the root hair MYB (Lee and Schiefelbein.

1999) and MYB5 is the outer seed coat MYB (Gonzalez et al 2009). The bHLHs are pleiotropic for all cell fates yet they appear to specialize and function better in some pathways than others. The major bHLHs are Glabra3 (GL3)(Payne et al. 2000), Enhancer of Glabra3 (EGL3)(Zhang et al. 2003) and Transparent Testa 8 (TT8)(Nesi et al. 2000). The use of different MYBs for each pathway was thought to be the mechanism allowing the same transcription complex to differentially influence many different cell fates. However, during investigation of the various TTG1 dependent pathways, it was determined by RT-PCR, ChIP, and other methods that the complex appears to have the same transcriptional targets repeating across the various cell fate pathways, namely GL2 and TTG2 (Morohashi et al. 2007; Zhao et al. 2008). This leads to the question of how a pair of proteins can regulate such different cell fates.

GL2 has been shown to be important for proper epidermal development. GL2 mutants have altered trichome elongation, possible reduction of trichome initiation, a gain of root hair initiation, and have altered seed coat cell morphology including a lack of seed coat mucilage (Fig1.2). GL2 mutants also have been described as having increased seed oil content (Shen et al. 2006). The GL2 protein is 747 amino acids long and contains four distinct domains; an acidic region, a Homeodomain (HD), a START domain, and a conserved carboxy end domain with no predicted structure. The acidic region has not been mentioned in the literature but is predicted by Uniprot. It's precise function is unknown.

The StAR-related lipid transfer (START) domain was originally found in mammals, but has since been identified in plants and other species. The START family in *Arabidopsis* includes 35 members of which 21 also contain a Homeodomain. These 21

are further divided into two classes; HD-Zip (5 members) and HD-ZLZ or HD-GL2 (17 members) (Schrick et al. 2004). The START domain consists of about 200 amino acids and forms a hydrophobic tunnel with proposed lid structure (Schrick et al. 2004; Alpy et al. 2005). In mammals, the START domain is also often coupled with other domains, but not to transcription factors. This domain is attached to proteins implicated in modulating signaling events, lipid metabolism, and lipid transfer or exchange between membranes or cellular compartments. This last appears to be the most common function. When not involved in lipid transfer the domain is predicted to function as a lipid-sensing domain (Alpy et al. 2005).

Currently no ligands have been identified for the plant START domain proteins however a number of the mammalian ligands have been identified. These include cholesterol, 25-hydroxycholesterol, phosphatidylethanolamine, ceramides, and phosphatidylcholine (Alpy et al. 2005). Based on sequence and predicted structure it is believed that some of the plant START domains will bind phosphatidylcholine, though no lipid ligand for the HD-GL2 group has been identified or predicted (Schrick et al. 2004).

One model for the HD/START proteins predicts that they can act in much the same way as mammalian steroid hormone receptors (Schrick et al. 2004) where binding of the lipid acts to facilitate or repress the function of the transcription factor. One of GL2's three identified transcriptional targets is a phospholipase (Ohashi et al. 2003). So it is conceivable that GL2's phospholipase target feeds back to modulate GL2's own functions via the START domain. With some mammalian hormone receptors, the hormone induces nuclear localization. There is anecdotal evidence that GL2 may

function similarly. In immunolocalization studies using GL2 antibodies, it was shown that GL2 was localized to the nucleus in trichomes, but was mostly cytoplasmic in subepidermal cells of young leaves (Szymanski et al. 1998). In the same study, GL1 was entirely localized to the nucleus. While lipid directed nuclear localization is possible, so are many other possibilities.

Another feature of the GL2 family is that the C-terminus of the protein is highly conserved across all higher plants (Fig 2; Schrick et al. 2004). This region contains approximately 250 to 300 amino acids and no predicted structure. High conservation strongly suggests a function, therefore I hypothesize that it is involved in protein binding. Using yeast 2-hybrid analysis, the C-terminus of the related HD-Zip group has been shown to interact with DORNROESCHEN (DRN), an ethylene response factor, via its AP2 domain (Chandler et al. 2005). Nothing is yet known about any partners for the HD-GL2 proteins.

The HD region of the HD-GL2 family has been determined to have a putative homeodomain binding site called the L1-box, taaatg(c/t)a (Abe et al. 2001). This can be found in the promoter of GL2's known targets (Ohashi et al. 2003; Tominaga-Wada et al. 2009). A search of other known proteins involved in the same pathways as GL2 has revealed that *EGL3, TT8, MYC1*, *MYB5*, *TT2*, and *TRY* contain L1-boxes but it is not known whether GL2 feeds back to regulate these TTG1 complex members or whether any HD-GL2 family member regulates any of these loci. A study examining the possible binding sites of the HD-GL2 family has shown that the most important region of the L1-box for binding is the TAAAT sequence (Nakamura et al. 2006).

GL2 has three known transcriptional targets involved in root hair development. The first target is phospholipaseD (PLDZ) which involved in promoting the root hair formation. GL2 has been shown to directly repress PLDZ (Ohashi et al. 2003). PLDZ has been predicted to be involved in vesicle trafficking, most likely for cell growth. The other known direct targets seem to be involved in production of cell wall polysaccharides. GL2 negatively regulates the cellulose synthase gene CESA5 and positively regulates XTH17 which is involved in constructing and organizing the cellulose/xyloglucan network in the cell wall. During the discovery of these two targets a new phenotype for GL2 was identified. GL2 mutants have increased cellulose production.

GL2 in the seed and embryo has been predicted to affect the expression of WER and mucilage modified 4 (MUM4). WER, the TTG1 complex member MYB which regulates root hairs, expression is increased in GL2 mutant embryos. WER does have an L1-box though it is 200 bp downstream of the stop codon (Costa and Dolan 2003). MUM4, a gene responsible for the production of seed coat mucilage, expression requires GL2 in the seed coat (Western et al. 2004). MUM4 however does not have a L1-box.

TTG2 is a WRKY domain containing transcription factor consisting of 429 amino acids. It is required for proper trichome and seed coat morphology, and mucilage and tannin production in the seed coat (Fig 1.2; Johnson et al. 2002). The WRKY domain is a DNA binding domain of about 60 amino acids containing the conserved WRKY motif along with a novel zinc finger motif (Eulgem et al. 2000). WRKY proteins show high affinity for the W box, a DNA element found in the targets of WRKY transcription factors and defined as (T)(T)TGAC(T/C) (Eulgem et al. 2000). As a family, WRKY

genes tend to mediate biotic and abiotic stress responses. TTG2 is biologically exceptional in that it functions in development.

TTG2 mutants have a trichome phenotype of reduced initiation (fewer trichomes), reduced branching, and wall thickening. Many of the trichomes are spikes that appear hyper elongated. GL2 mutants, in addition to "nub" trichomes that lack vertical elongation, produces some spikes and single branched trichomes, though they are generally shorter than the ttg2 mutant trichomes. The comparison between gl2 and ttg2 mutants shows the specialization of each in the development of trichomes; GL2 is key for elongation while TTG2 is important for branching and other maturation processes. TTG2 like GL2 is expressed in the non-hair files of roots; however, it currently has no known role in their development.

TTG2 has many roles in the seed. In addition to being important for seed coat morphology, mucilage production and testa pigmentation TTG2 mutants have reduced seed size compared to wild type. It is thought that TTG2 controls seed size by regulating integument cell elongation (Garcia et al. 2005).

Unlike GL2, TTG2 has no transcriptional targets identified yet. In fact while there is almost nothing known about how GL2 functions, there is even less known about TTG2. Most of the information about TTG2 function consists of negative results. For example, althought TTG2 is required for tannin production, it has been shown not to regulate BANYULS, an early enzymatic step in tannin biosynthesis (Debeaujonor et al. 2003). Similarly, while TTG2 is required for seed coat mucilage production, it does not regulate MUM4, an enzymatic step required in mucilage biosynthesis (Western et al. 2004). TTG2 has been suggested to have a role in the expression of GL2, however the

evidence consists of a single experiment in the root with an overexpressed TTG2 protein fused to a strong repressing domain (Ishida et al. 2007). While this finding may be true, nothing is known about GL2 or TTG2 that suggests that this could be the case. It has been shown in numerous publications that GL2 expression does not change in any tissue in ttg2 mutants, including Ishida et al. (2007).

This work described here outlines progress made towards verifying the TTG1 complex transcription control of GL2 and TTG2, characterization of GL2 protein function, and identification of GL2 and TTG2 transcriptional targets.

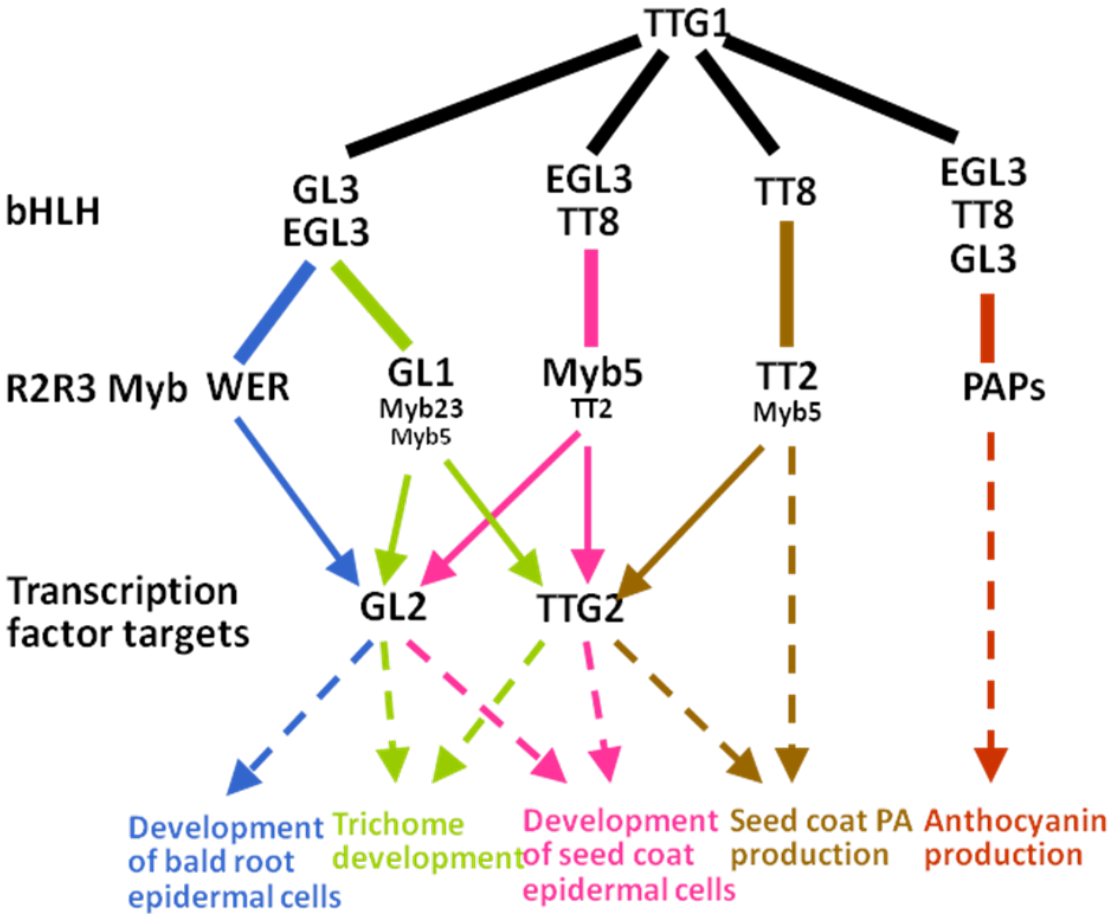

Fig 1.1. TTG1 Combinatorial Transcription Complex.

Solid lines represent protein-protein interaction, solid arrows represent direct transcriptional activation and dashed arrows represents control of the process. Redraw with permission from Gonzalez et al (2009).

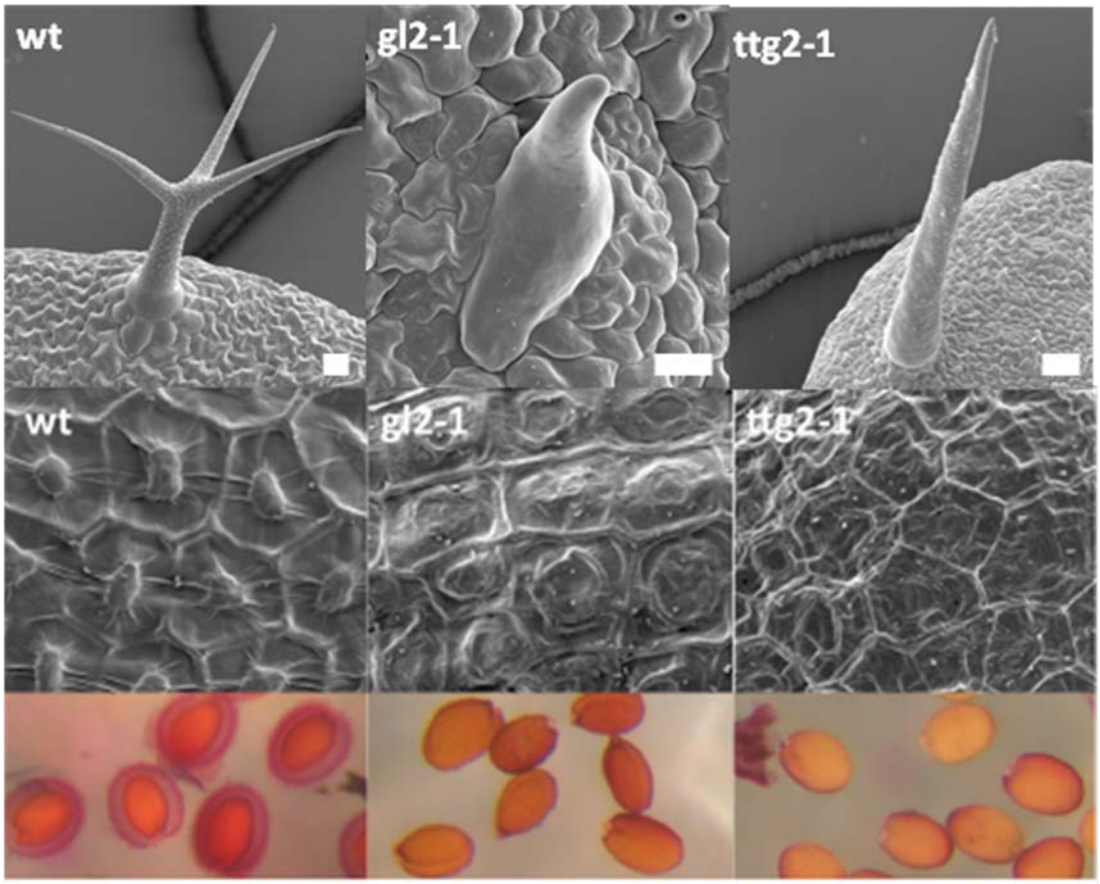

Fig 1.2. *gl2-1* and *ttg2-1* trichome and seed phenotypes.

SEM images of trichomes and seed coats of wild type Col, *gl2-1* and *ttg2-1* mutants. Size bars in the trichome images represent 20 μm. Ruthenium red stained seeds of wild type Col, *gl2-1* and *ttg2-1*. Gl2-1 trichomes stall at the elongation stage while ttg2-1 has defects in branching and the other late stages of maturation, while both have similar defects in seed coat morphology and mucilage production.

Chapter 2: The TTG1-bHLH-MYB complex controls trichome cell fate and patterning through direct targeting of regulatory loci

INTRODUCTION

This chapter consists of work done in collaboration with Mingzhe Zhao of Alan Lloyd's lab, Kengo Morohashi and Erich Grotewold at The Ohio State University. It was published in Development in 2008 (Zhao et al. 2008). I contributed the localization and movement studies.

Studies on the regulation of cell fate and function on the plant epidermis continue to provide important insights into how plant cells are organized, how patterning develops, and how developmental and biochemical pathways interact. Trichome initiation in the model plant *Arabidopsis thaliana* has been an important model for understanding cell fate and patterning. Trichomes (leaf hairs) are large, branched, single cells which initiate and develop on young leaves in a regular spacing pattern (Larkin et al., 1997; Marks, 1997; Hulskamp and Schnittger, 1998; Hulskamp et al., 1999). Trichome patterning is not random or dependent on other cell types or position on the leaf, but is thought to be generated *de novo* by intercellular communication (Larkin et al., 1996; Schnittger et al., 1999). The model assumes that inhibitors, activated by self-enhanced activators, can move between cells to mediate competition between equivalent cells, resulting in the pattern formation (Larkin et al., 2003; Pesch and Hulskamp, 2004).

Years of genetic and molecular studies have enabled the identification of components of this trichome patterning machinery. Three classes of interacting regulators including the

R2R3-MYB transcription factor, GLABRA1 (GL1) (Oppenheimer et al., 1991), the basic helix-loop-helix (bHLH) proteins, GLABRA3 (GL3) and ENHANCER OF GLABRA3 (EGL3) (Payne et al., 2000; Zhang et al., 2003) and the WD40 repeat protein, TRANSPARENT TESTA GLABRA1 (TTG1) (Walker et al., 1999) are postulated to form a combinatorial regulatory complex. Evidence comes from yeast two-hybrid studies showing that TTG1 and GL1 physically interact with GL3/EGL3 but not with each other (Payne et al., 2000; Zhang et al., 2003). *GLABRA2 (GL2)* is a direct target of GL3 and EGL3 (Morohashi et al., 2007) and *TRANSPARENT TESTA GLABRA2 (TTG2)* is directly regulated by GL1 (Ishida et al., 2007). This activation is believed to be through the formation of TTG1-GL3-GL1 and TTG1-EGL3-GL1 (TTG1-bHLH-GL1) regulatory complexes (Szymanski et al., 1998), thereby regulating trichome cell fate. *GL2*, a homeodomain (HD-Zip) and *TTG2,* a WRKY transcription factor, are required for normal trichome development (Rerie et al., 1994; Johnson et al., 2002). Some levels of GL2 overexpression can result in trichome clusters indicating that this HD-Zip may function in the regulation of trichome spacing (Ohashi et al., 2002).

To date, a group of at least four homologous single MYB proteins, TRIPTYCHON (TRY) (Schellmann et al., 2002), CAPRICE (CPC) (Wada et al., 1997) and ENHANCER OF TRY and CPC1 and 2 (ETC1 and 2) (Kirik et al., 2004a; Kirik et al., 2004b), have been identified as negative regulators of trichome initiation and patterning. The *try cpc* double and the *try cpc etc1* triple mutants (Kirik et al., 2004a; Schellmann et al., 2002) display a greatly enhanced “clustered-trichome” phenotype, indicating that lateral inhibition is disrupted. These inhibitory proteins contain no recognizable transcription activation domain. Therefore, they could work as negative

transcriptional regulators. Protein interaction analysis in yeast has suggested that TRY or CPC would interrupt the functionality of the "activating" TTG1-bHLH-GL1 complex by competitive interaction with the bHLH (Esch et al., 2003; Zhang et al., 2003). Additionally, the individual members of this inhibitory protein family may function differently. There is evidence that TRY might be more important in short-range inhibition while CPC and particularly ETC1 may be important for long-range inhibition (Schellmann et al., 2002; Kirik et al., 2004a).

As described above, the identification of these positive and negative trichome regulators has laid an excellent foundation for understanding trichome patterning. However, a large amount of the data elucidating the molecular mechanism of these regulators is either indirect or obtained from another similar pathway - root hair patterning. For instance, evidence for the existence of the TTG1-bHLH-MYB complex is based entirely on protein interaction studies in yeast (Payne et al., 2000; Zhang et al., 2003; Zimmermann et al., 2004). Furthermore, the only evidence demonstrating the ability of a single MYB inhibitor to move between cells is that CPC-GFP fusion protein is detected both in the trichoblasts and in the atrichoblasts in roots when its transcript is only found in hairless cells (Wada et al., 2002). More importantly, the regulatory events triggered by the TTG1-bHLH-MYB active complex mostly remain unknown. The expression of *CPC* in the root epidermis is *GL3/EGL3* dependent (Bernhardt et al., 2005) and directly regulated by the MYB WEREWOLF (WER) (Lee and Schiefelbein, 2002; Koshino-Kimura et al., 2005; Ryu et al., 2005), a GL1 equivalent protein in root hair patterning (Lee and Schiefelbein, 2001).

In recent work, we have shown that *GL2, CPC* and *ETC1* are directly activated by GL3 and this targeting is GL1 dependent (Morohashi et al., 2007). The work presented here is aimed at further testing and refining details of the trichome development model under the control of the TTG1-bHLH-MYB complex. Here we show that the trichome activators, *GL2 and TTG2,* and repressors, *CPC and ETC1*, are major transcriptional targets for the complex. In addition, we also demonstrate the existence of the TTG1-bHLH-MYB complex in plants and show that loss of TTG1 or GL1 disrupts the distribution of GL3. Furthermore, we demonstrate that the CPC protein moves in the leaf epidermis while none of the activators tested move. These results support major aspects of the model and also add novel perspectives to the current model for trichome patterning.

RESULTS

TTG1 is expressed ubiquitously in Arabidopsis leaves

The transcription of TTG1 is detected in all major organs of *Arabidopsis* (Walker et al., 1999). To study the expression of the TTG1 protein during the process of trichome initiation and patterning, we examined YFP fluorescence of a YFP-TTG1 fusion protein under the control of the native *TTG1* promoter in the *ttg1* mutant background *(ttg1*/pTTG1::YFP-TTG1). The transgenic *ttg1*/pTTG1::YFP-TTG1 plants showed wild type trichome formation (Fig 2.1), as well as normal anthocyanin production and seed coat pigment and differentiation (not shown), indicating that the translational YFP-TTG1 fusion was functional. At early stages strong YFP signal is detected in the nuclei of trichome initials and of all pavement cells, with a much weaker YFP signal in the

cytoplasm of these cells (Fig 2.1A). Fig 2.1B shows that the TTG1 protein is present in all epidermal cells and trichomes at all developmental stages. This ubiquitous and persistent expression pattern was further confirmed by a close-up view of the YFP expression in stage-5 trichomes (Szymanski et al., 1998) and their surrounding epidermal cells (Fig 2.1C). TTG1 protein appears to be expressed at all stages of leaf and trichome development.

TTG1 regulates GL3 target genes

GL3 has been reported to directly target genes that regulate trichome development, both trichome activators and repressors (Morohashi et al., 2007). To better define the trichome genes regulated by the TTG1-bHLH-MYB regulatory complex, the expression changes of previously identified GL3 targets including *GL3, GL2, ETC1* and *CPC* (Morohashi et al., 2007) were investigated by quantitative PCR (Q-PCR) in DEX-treated *ttg1*/p35S::TTG1-GR plants, in the presence or absence of CHX. The TTG1-GR fusion complements *ttg1* mutants only with the addition of DEX. Simultaneous treatment with DEX and CHX blocks *de novo* protein production and allows only the direct targets to be transcribed (Sablowski and Meyerowitz, 1998). This same TTG1-GR line has been used to show that the bHLH, TT8, was directly activated by TTG1 in siliques (Baudry et al., 2006). In other work, we have shown that this fusion provides DEX dependent activation of the late anthocyanin structural genes (Gonzalez et al., 2008).

As shown in Fig 2.2A, *GL2*, *CPC* and *ETC1* were up-regulated in response to the 4-hour induction by DEX, while the expression of *GL3*, *TRY* and *ETC2* did not change. This experiment was repeated with a DEX plus CHX treatment. *GL2*, *CPC*, and *ETC1* again

were up-regulated, but to a lower level (Fig 2.2A). A two-sided t-test indicates that these induction levels are significantly greater that uninduced levels ($P < 0.05$). To confirm that these expression results are due to direct activation, chromatin immunoprecipitation (ChIP) experiments were performed with *ttg1*/pTTG1::YFP-TTG1 plants, using antibodies against GFP which cross-react with YFP. Similar to what was previously described for GL3 (Morohashi et al., 2007), YFP-TTG1 was recruited to the promoters of *GL2*, *CPC* and *ETC1* in vivo (Fig 2.2B). These results show that *GL2*, *CPC* and *ETC1* are immediate direct targets of TTG1, indicating that TTG1 and GL3 share many of the same targets.

TTG2 is an immediate direct target of TTG1 and GL3

Genetic data show that the expression of *TTG2* requires *TTG1* (Johnson et al., 2002), suggesting that TTG1 and GL3 directly control *TTG2* expression in vivo. We analyzed *ttg1*/p35S::TTG1-GR and *gl3 egl3*/p35S::GL3-GR transgenic seedlings for expression changes in *TTG2* after DEX induction. Four hour DEX induction of TTG1-GR and GL3-GR resulted in the up-regulation of *TTG2* (Fig 2.2A). Inclusion of DEX and CHX also resulted in the significant induction ($P < 0.05$, two-sided t-test) of *TTG2* and identified it as a direct target of both TTG1 and GL3. This finding is confirmed by ChIP results unequivocally demonstrating that TTG1 binds to the promoter of *TTG2* in vivo (Fig 2.2B).

GL1 participates in the regulation of *GL2, TTG2, CPC* and *ETC1*

It was previously shown that WER binds the *CPC* promoter in vitro (Koshino-Kimura et al., 2005; Ryu et al., 2005) and it has been recently reported that GL1 directly

regulates *TTG2* (Ishida et al., 2007). We have shown that GL3 regulates and binds the promoters of *GL2*, *CPC*, *ETC1* and *GL3* in vivo. The binding of GL3 to the *CPC* and *GL2* promoters is dependent on the presence of GL1 while binding to its own promoter is not (Morohashi et al., 2007). We used *gl1*/pGL1::GL1-YFP-cMYC plants to perform ChIP experiments to investigate the in vivo binding of GL1 to the promoters of these known GL3 targets. GL1 was found to bind to the promoters of *CPC*, *ETC1* and *GL2*, as well as that of *TTG2*, but not of *GL3* (Fig 2.2B). These results suggest that GL1 participates with GL3 in the regulation of *GL2*, *CPC*, *ETC1* and *TTG2*, but not in the auto-regulation of *GL3*. Taken together with the finding that *GL2*, *TTG2*, *CPC* and *ETC1* are direct transcriptional targets of both GL3 and TTG1, while *GL3* is only regulated by GL3, it is most likely that *GL2*, *TTG2*, *CPC* and *ETC1* are activated by a complex containing TTG1, GL3 and GL1.

It is interesting that the QRT-PCR analyses with cycloheximide seem to reveal additional, non-TTG1-dependent regulatory effects with GL2 being activated and TTG2 being repressed by other factors. In these experiments, we only conclude that a gene is a direct target if the RT-PCR and the ChIP experiments are in agreement.

TTG1 interacts with GL3 and GL1 in vivo

The gene expression studies presented above support the hypothesis that TTG1 participates in a TTG1-bHLH-MYB activation complex but does not directly demonstrate that TTG1 and GL1 co-exist in a complex. To detect this complex in vivo, we performed co-precipitation assays to test whether TTG1 interacts with GL3. The *TTG1::TTG1-cMYC* and *35S::HA-GL3-6His* fusions are functional in promoting trichome

differentiation in *ttg1* and *gl3 egl3* mutants respectively. As shown in Fig 2.3A, the TTG1-cMYC fusion was detected in the input protein extractions of plants containing this construct (lanes 1 and 3) using an anti-cMYC monoclonal antibody. However, when His-select Ni columns were used to pull down the HA-GL3-6His fusion protein from these extracts, TTG1-cMYC was detected only in the line containing both fusion proteins (Fig 2.3A lane 6), demonstrating that TTG1 interacts with GL3 in vivo.

Using the same approach, we also tested for the interaction between TTG1 and GL1 in vivo. Strikingly, TTG1-cMYC was pulled down by the His-select Ni columns only when it was co-expressed with GL1-YFP-6His (Fig 2.3B lane 6), while TTG1-cMYC was not detected in the samples processed from *ttg1*/pTTG1::TTG1-cMYC or *gl1*/pGL1:: GL1-YFP-6His (Fig 2.3B lane 4, 5) demonstrating that TTG1 interacts with GL1 in vivo. These results do not indicate that TTG1 directly touches GL1 and when combined with yeast 2-hybrid analysis (Payne et al., 2000; Zhang et al., 2003), these results indicate that TTG1 and GL1 interact by both binding to GL3 or EGL3 as intermediates.

Loss of TTG1 and GL1 disrupts the nuclear distribution of GL3

Experiments were performed to test whether TTG1 and GL1 affect the GL3 protein distribution pattern in the leaf epidermis. A functional *GL3::GL3-YFP* fusion (Bernhardt et al., 2005) was examined in the wild type plant. We detected GL3-YFP signal restricted to the nuclei of trichome cells with an evenly distributed fluorescence pattern (Fig 2.6E). When *GL3::GL3-YFP* was introduced into the *ttg1* mutant background, no obvious changes in GL3's partitioning to the nucleus was observed. However, the GL3-YFP protein was unevenly distributed into speckles in the nuclei of

epidermal cells (Fig 2.4A). In contrast, epidermal cells of the *ttg1* mutant showed evenly distributed GL1-YFP with only a couple of speckles in the nucleus (Fig 2.4B). These results suggest that TTG1 is required for the proper subnuclear distribution of GL3. Although it is difficult to quantitatively compare these images, it does not appear that loss of TTG1 affects the stability of the GL3-YFP fusion.

To test whether mutations in GL1 might affect the distribution of GL3, we examined the subcellular localization of GL3-YFP in the *gl1* mutant. When *GL3::GL3-YFP* was expressed in the *gl1* mutant, GL3 still partitioned to the nucleus. However, just like the *ttg1* mutant, GL3 formed speckles in the nuclei of leaf epidermal cells (Fig 2.4C). In the roots of the same transgenic plant, where GL1 function is replaced by WER, GL3-YFP showed wild type patterning with no speckles (Fig 2.4D). These results suggest that GL1 is specifically required for the normal distribution of GL3 within the nuclei of *Arabidopsis* leaf cells.

Taken together, our studies on *in vivo* protein interactions and the subcellular localization of fluorescent fusion proteins show that TTG1, GL3 and GL1 form a nuclear complex in vivo. Moreover, the loss of TTG1 or GL1 leads to an abnormal speckled distribution of GL3, a key complex member.

CPC moves in leaf epidermal cells

It has been shown that GL3 and CPC traffic from cell to cell in the *Arabidopsis* root epidermis to specify near neighbor cell fate (Wada et al., 2002; Bernhardt et al., 2005). We hypothesized that similar movements might be required during trichome patterning events. YFP fusions to TTG1, GL3, GL1, CPC and GL2 were used to examine

whether any of these proteins could move from cell to cell in the leaf epidermis. The fusion genes were introduced into developing leaf tissue by microprojectile bombardment and were scored after overnight expression. We also bombarded a *35S::GUS* reporter and we did not detect any area with clusters of transformed GUS-expressing cells, indicating that the probability of bombarding adjacent cells is very low (data not shown).

We repetitively observed extensive trafficking of the YFP-CPC fusion into adjacent cells as evidenced by cytoplasmic and nuclear YFP signal (CPC moved in 32 of 76 bombardment events), generating clusters of up to 15 fluorescent cells in the *Arabidopsis* leaf epidermis (Figs. 5D and S1). In contrast, we did not observe the same fluorescent pattern with any of the other fusion proteins, which were expressed in isolated single cells (Fig 2.5); at least 50 bombardment events were observed with each gene. These results show that CPC can move in the leaf epidermis, but that GL3 does not. Our results, showing that CPC but not GL3 moves in the leaf epidermis, contrast with previous findings that they both move in roots (Bernhardt et al., 2005; Wada et al., 2002). This probably reflects that fact that trichome patterning and root hair patterning are not regulated by the same mechanisms, although they largely share the same hierarchy of regulatory genes. In roots, GL3 must move from the hair cell files, where it is transcribed, to the hairless files where it functions. In leaves it is transcribed in the trichome initials, where it functions, and so is not required to move. It is possible that there is some developmental control of intercellular movement, however, we observed CPC movement no matter where on the leaf we bombarded while GL3 never moved. We note that we were not able to successfully bombard the very youngest and smallest cells on the leaf epidermis.

GL3 and EGL3 have overlapping but distinct expression patterns

Our previous studies showed that *GL3* and *EGL3* are partially redundant in regulating trichome initiation. The *gl3 egl3* double mutant is completely glabrous (Zhang et al., 2003). However, single *gl3* mutants show a much more severe reduction in trichome initiation and branching than *egl3* (Zhang et al., 2003) and we have shown that *GL3* but not *EGL3* participates in an auto-regulatory loop (Morohashi et al., 2007). In order to begin to characterize the functional differences between GL3 and EGL3, we carefully examined the *Promoter::GUS* expression and protein accumulation patterns of *GL3* and *EGL3* during trichome development in wild type plants.

Maximum GUS activity was observed in young leaf primordia for both *GL3* and *EGL3* (Fig 2.6A, C). In young developing leaves, GL3::GUS activity is observed especially in the region close to the basal edge of the leaf (Fig 2.6A). In the same age leaves, high EGL3::GUS activity is observed in the basal one third of the leaf and is not restricted to the edge (Fig 2.6C). In both lines, maturing and mature trichomes show significantly higher levels of GUS activity than surrounding epidermal cells (Fig 2.6A, C). In more mature leaves, strong GL3::GUS activity becomes restricted to trichomes (Fig 2.6B), while EGL3::GUS persists at low levels in pavement cells as well as in trichomes (Fig 2.6D). Compared to *GL3*, *EGL3* exhibits a more widely distributed transcription pattern with higher GUS activity in the epidermal pavement cells that GL3 and lower GUS activity in trichomes than GL3. High EGL3::GUS activity is also observed in the petioles of leaves while GL3::GUS is not. Taken together, GL3 and EGL3 show overlapping, yet distinct transcription patterns during trichome development.

The *GL3::GL3-YFP* and *EGL3::EGL3-YFP* fusions were constructed and shown to be fully functional by rescuing *gl3 egl3* mutants (not shown). The analysis of the YFP fluorescence profiles of representative wild type *GL3::GL3-YFP* and *EGL3::EGL3-YFP* containing transgenic plants shows that the protein expression profiles of GL3 and EGL3 generally match well with their transcription patterns respectively with some notable differences.

In the basal region of the developing leaf, where trichomes continue to initiate, strong GL3-YFP signal was detected in the nuclei of unbranched trichome initials while only a very weak GL3-YFP signal was occasionally detected in the neighboring non-trichome cells (Fig 2.6E arrows, F). As a trichome matures, the level of GL3-YFP intensity keeps decreasing until it completely disappears (not shown). Like GL3-YFP, EGL3-YFP was also found to increase in the nuclei of trichome initials in the leaf basal region but not as high as GL3. However, EGL3-YFP was also detected in the nontrichome cells throughout the epidermal layer of a developing leaf (Fig 2.6G, H).

A comparison of patterns of *GL3::GUS* and *GL3::GL3-YFP* reveals a difference between the transcription pattern and the protein expression pattern of GL3. Significant GL3::GUS activity was observed in the epidermal cells that neighbor young trichomes where GL3 protein is absent (compare Fig 2.6A with E). Taken together with the finding that EGL3 gene is expressed and the EGL3 protein accumulates in both trichome and non-trichome cells, these data imply that EGL3 functions within the non-trichome cell in the maintenance of the non-trichome cell fate, while GL3 does not.

MATERIAL AND METHODS

Plasmids

Plasmid descriptions are below. Details of plasmid constructions will be provided on request. Sequences of primers used are provided in Table 2.S1.

pTTG1::YFP-TTG1 and *pTTG1::TTG1-cMYC* contain a *TTG1* genomic fragment, including 1 kb of 5' and 3' regulatory sequences, and either the *YFP* coding region without a stop codon inserted in frame with the *TTG1* start codon, or 5 copies of the *cMYC* epitope inserted in frame with an altered *TTG1* stop codon.

pGL1::GL1-YFP-cMYC and pGL1::GL1-YFP-6His contain a *GL1* genomic fragment, including 1.45 kb of 5' and 1 kb of 3' regulatory sequences, and either a *YFP-cMYC (5XcMYC)* or *YFP-6His* fusion inserted in frame with an altered *GL1* stop codon.

p35S::HA-GL3-6His contains the CaMV35S promoter driving the *GL3* genomic coding region with both the *HA* epitope in frame with the *GL3* start codon and the *6His* epitope in frame with an altered *GL3* stop codon.

p35S::GL3-YFP, p35S::GL1-YFP, p35S::GL2-YFP and *p35S::YFP-CPC* contain the entire *GL3* or *GL1* genomic coding regions or the *GL2* coding cDNA, with the stop codons removed, or the entire *CPC* cDNA cloned into appropriate CaMV35S-YFP fusion cassette vectors.

pEGL3::EGL3-YFP contains a 6 kb *EGL3* genomic fragment containing 3 kb upstream of the start with a deleted stop codon cloned in frame to *YFP*.

Plant materials and growth conditions

L*er*/pGL3::GUS and L*er*/pEGL3::GUS were described previously (Zhang et al., 2003). *gl3 egl3*/p35S::GL3-GR was described previously (Morohashi et al., 2007).

ttg1/p35S::TTG1-GR seeds (Baudry et al., 2006) were generously provided by Dr. Loic Lepiniec. *gl3-2*/pGL3::GL3-YFP was previously described (Bernhardt et al., 2005). To generate *gl1*/pGL3::GL3-YFP and *ttg1*/pGL1::GL1-YFP-cMYC, *gl3-2*/pGL3::GL3-YFP was crossed to *gl1* and *gl1*/pGL1::GL1-YFP-cMYC was crossed to *ttg1* respectively. Plants expressing both *TTG1::TTG1-cMYC* and *GL1::GL1-YFP-6His* fusions were created by crossing *gl1*/pGL1::GL1-YFP-6His to *ttg1*/pTTG1::TTG1-cMYC. F2 plants were confirmed by YFP fluorescence microscopy and western blots probed with an anti-cMYC antibody. Lines expressing both *TTG1::TTG1-cMYC* and *35S::HA-GL3-6His* fusions were created by transforming *ttg1*/ pTTG1::TTG1-cMYC plants with p35S::HA-GL3-6His. Transformants were identified by kanamycin and BASTA double resistance. All transgenic plants were created by floral dip transformation. Standard plant crosses were done with two homozygotes and the F1 were selfed to identify proper progeny. *Arabidopsis* plants were grown on soil at 21°C in continuous white light.

Gene expression analyses

Seedlings were grown on MS media containing 3% sucrose at 21°C in continuous white light. Four-day-old seedlings were treated with 20 μM dexamethasone (DEX) or mock-treated with 0.001% ethanol for four hours, washed with water and frozen in liquid nitrogen. 100 □M cycloheximide (CHX) treatment was used when appropriate. Total RNA was prepared according to Morohashi et al. (2007). 4 μg of RNA was used in 20 μl reverse transcription reactions containing 250 nM Actin and target gene specific reverse primers. Parallel 25 μl PCR reactions were prepared using cDNA reactions as templates with half volume of 2X SuperPower Syber mixture (ABI) and run on a

spectrofluorometric thermal cycler (ABI 7900HT). For each target, five PCR reactions containing 400 nM primers and 3 μl first strand target gene cDNA as template were performed along side four actin control PCR reactions containing 200 nM Actin primers and 1 μl first strand Actin cDNA. The comparative cycle threshold method was used to analyze the results (User Bulletin 2, ABI PRISM Sequence Detection System). Each experiment was performed twice for each target with consistent results. Results of representative experiments are presented.

Chromatin Immunoprecipitation (ChIP) Experiments

ChIP experiments were performed as described previously (Morohasi et al., 2007).

Microscopy

The histochemical analyses of *Promoter::GUS* reporter genes were performed with at least 5 seedlings for each strain essentially as described (Masucci et al., 1996). Imaging of YFP fusions was performed on a Leica SP2 AOBS confocal laser scanning microscope with excitation (514 nm) and emissions (530–600 nm for YFP and 675-800 for chlorophyll). Collected images were processed for maxium intensity projection.

Microprojectile Bombardment

Tungsten particles (1.5 mg) were coated with approximately 5 μg of each plasmid DNA as directed by the manufacturer's instructions (Bio-Rad). Young leaves from *gl3 egl3* double mutant plants were excised, placed on MS plates and bombarded at 1100 psi with a flight distance of 15 cm using a Bio-Rad PDS-1000. Bombarded leaves were placed overnight under white light and imaged on the confocal microscope. At least three

independent bombardment experiments were performed with each construct with multiple bombardment events in each experiment so that over 50 events were observed for each construct.

Co-precipitation experiments

Three-week-old *Arabidopsis* green tissue was ground into fine powder in liquid nitrogen. Protein extract was prepared by thorough mixing of 0.1 g powder with 1 ml ice-cold buffer A (50 mM Tris, 100 mM NaCl, 10 mM $MgCl_2$, 10% glycerol, 1 mM DTT, 1% Triton-X100, 1 mM PMSF, 1 μg/ml each of (Leupeptin, Antipain, Pepstatin A, Aprotinin), 5 mM imidazole, pH 7.3) in a 1.5 ml eppendorf tube. The mixture was centrifuged twice at 13,000 rpm for 10 minutes, and the supernatant was used as input extract. 0.9 ml of the input extract was applied to a pre-equilibrated His-select column (with buffer A), washed (with buffer A containing 45 mM imidazole) and eluted (with buffer A containing 300 mM imidazole) as directed by the manufacturer (Sigma). The elution was concentrated with Microcon Y-M30 filter (Millipore). Input extracts and concentrated eluates were mixed with loading buffer to final volume of 100 μl and boiled for 5 minutes prior to loading the SDS-PAGE gel (Bio-Rad). Two μl of input and 5 μl of elution loading samples were used for western blots, which were probed by anti-cMYC monoclonal antibody 9E10 (Santa Cruz Biotechnology) and visualized by Western Lightning Chemiluminescence Reagents (Amersham Biosciences).

DISCUSSION

The *Arabidopsis TTG1* locus encodes a WD40 protein containing four WD40 repeat motifs but no recognizable nuclear localization signal, DNA binding motif or

transcriptional activation domain (Walker et al., 1999). A common function of WD40 repeat motifs is to facilitate protein-protein interactions. A preponderance of indirect evidence indicated that TTG1 interacts with bHLH proteins (GL3, EGL3 and TT8) in regulating all TTG1-dependent development pathways. The evidence includes: (1) *gl3 egl3 tt8* triple mutant phenocopies *ttg1* mutant and, (2) TTG1 physically interacts with GL3, EGL3 and TT8 in the yeast-two hybrid system (Payne et al., 2000; Zhang et al., 2003; Baudry et al., 2004). TTG1 may form a complex with bHLH proteins for nuclear import or retention and/or act as a transcriptional co-regulator. Prior to the present study, it was also possible that TTG1 was located only in the cytoplasm possibly as a signal transduction component to regulate bHLH proteins. A cytoplasmic location would be in agreement with the reported location of AN11 (de Vetten et al., 1997), a petunia WD40 protein which is highly similar to TTG1 and complements the *ttg1* mutation (not shown). In this paper, we report that TTG1 is preferentially localized in the nucleus in the *Arabidopsis* leaf epidermis (Fig 2.1), with apparently a lower, yet significant, amount of TTG1 in the cytoplasm. This result indicates that TTG1 could function both as a transcriptional co-regulator in the nucleus and as a protein-interacting factor in the cytoplasm.

The TTG1-bHLH-MYB regulatory complex

Although we have demonstrated that TTG1 interacts with GL3 in vivo (Fig 2.3), the biological significance of the TTG1-bHLH interaction still remains to be elucidated. Our previous genetic data (Zhang et al., 2003), together with the results discussed in this paper, favors the possibility that TTG1 functions as a transcription co-regulator. TTG1

may modify, stabilize or in some other fashion positively affect GL3/EGL3 in their capacity to activate the transcription of downstream target genes. Our work on the regulation of the anthocyanin pathway shows that GL3 and TTG1 regulate the same set of anthocyanin biosynthetic target genes (Gonzalez et al., 2008). It would not be surprising that TTG1 and GL3 regulate the same target genes in the trichome development pathway. Our results using a TTG1-GR inducible system show that *GL2, CPC* and *ETC1* are also direct targets of TTG1, because the transcription of these genes increased significantly in response to TTG1-GR induction even in the absence of *de novo* protein synthesis. We have also identified *TTG2* as an immediate direct target of both TTG1 and GL3 (Fig 2.2), which is consistent with the finding that TTG2 is directly regulated by GL1 (Ishida et al., 2007). These data show that TTG1 largely regulates the transcription of the same regulatory loci as GL3 during trichome cell fate specification. It also supports the notion that TTG1 regulates the trichome pathway through affecting the activation capacity of bHLH proteins.

Interestingly, we failed to detect any changes in *GL3* expression after TTG1-GR induction, as opposed to the finding that *GL3* is repressed by GL3-GR (Morohashi et al., 2007). It has been reported that GL3 binds to and activates *GL2, CPC* and *ETC1* in a GL1-dependent manner but the GL3 self-repression is GL1-independent (Morohashi et al., 2007). In our ChIP experiments with *gl1*/pGL1::GL1-YFP-cMYC, we detected the in vivo recruitment of GL1 to the *GL2, TTG2, CPC* and *ETC1* promoters but not to the promoter of *GL3* (Fig 2.2B). These data suggest that the GL1 DNA-binding activity is required for the TTG1-bHLH complex to select target genes and that GL3 self-repression may be both GL1 and TTG1 independent. Additionally, the in vivo interaction between

TTG1 and GL1 (Fig 2.3B) fits perfectly with the model that TTG1, bHLH and R2R3-MYB proteins form a TTG1-bHLH-MYB regulatory complex in vivo. The TTG1-bHLH-MYB complex seems to only activate the transcription of downstream targets but not the transcription of bHLH or R2R3-MYB proteins in the trichome pathway.

We could not detect changes in the expression of *TRY* or *ETC2* by the induction of *gl3 egl3*/p35S::GL3-GR (Morohashi et al., 2007) or *ttg1*/p35S::TTG1GR (Fig 2.2A). These results demonstrate that although *TRY* and *ETC2* are largely redundant with *CPC* and *ETC1*, they are regulated differently, perhaps by GL2 for example, which is consistent with their different levels of expression in different tissues (Kirik et al., 2004b).

How does TTG1 function?

GL3 transcripts can be easily detected in the *ttg1* and *gl1* mutants (Payne et al., 2000) indicating that they are not required for GL3's transcription. We wanted to determine whether TTG1 might regulate the subcellular localization of GL3. In the *ttg1* mutant, we found that the GL3-YFP protein was still located entirely in the nucleus. Surprisingly, however, the loss of TTG1 caused GL3 to be abnormally distributed within the nucleus of leaf epidermal cells. GL3 protein forms unevenly distributed "speckles" (Fig 2. 4A). In contrast, the nuclear distribution pattern of GL1-YFP-cMYC in *ttg1* is very similar to the wild type pattern - a more or less even nuclear distribution. One or two GL1 speckles were found in a single nucleus (Fig 2.4C). These results suggest that functional TTG1 protein is required for the appropriate bHLH distribution in the nucleus but is largely not necessary for GL1 distribution.

In *gl1* mutants, we detected a similar but even more severely speckled GL3-YFP distribution, specifically in the leaf epidermis (Fig 2.4C, D). GL3 forms fewer but more clearly isolated nuclear speckles in *gl1* than in *ttg1* (compare Fig 2.4A and 2.4C). We previously showed that in a *gl1* mutant, GL3 is no longer recruited to the promoter of its major trichome targets, *GL2* and *CPC* (Morohashi et al., 2007). Taken together, we conclude that GL1 is responsible for GL3 or the TTG1-bHLH complex tethering to the promoters of specific downstream targets, and TTG1 may function as a "helper" for the bHLH::GL1 interaction. Loss of proper DNA and/or protein interactions leads to aberrant bHLH distribution. It will be important for future studies to provide direct molecular or biochemical evidence to confirm that TTG1 facilitates the interactions between bHLH and MYB proteins.

Besides participating in the TTG1-bHLH-MYB regulatory complex, TTG1 may regulate trichome genes through other mechanisms. It was recently shown that TTG1 physically interacts with GEM, a protein that modulates cell division and represses the expression of *GL2* and *CPC* in *Arabidopsis* roots. Overexpression of *GEM* caused increased root hair and decreased leaf trichome densities (Caro et al., 2007). Overexpressed GEM is shown to bind to the promoters of *GL2* and *CPC*, and is associated with the acquisition and/or maintenance of histone H3K9me2 (typical of silent heterochromatic regions) at these two genes. These data imply that the interaction between TTG1 and GEM could prevent GEM from joining a complex which represses the expression of *GL2* and *CPC* or other trichome genes.

Trichome patterning

In theoretical models (Meinhardt, 1994; Meinhardt and Gierer, 2000), it is proposed that *de novo* patterning often requires the local self-enhancement of activators in combination with lateral inhibition by inhibitors. Based on this theory, a common model is proposed for the *Arabidopsis* trichome and root hair patterning, in which single MYB repressors (CPC and TRY) are thought to be able to move (faster than activators if activators can also move) into neighboring cells (Pesch and Hulskamp, 2004). In support of such a model is the fact that although CPC-GFP proteins are expressed in non-root hair cell files, the CPC-GFP protein is also detected in the neighboring root hair files (Wada et al., 2002). In this paper, our microprojectile bombardment experiment with *35S::YFP-CPC* directly demonstrates CPC's ability to move in the leaf epidermis for the first time (Fig 2.5), strongly supporting the current trichome patterning model from this perspective. YFP-CPC protein was detected in clusters of epidermal cells generally one cell, but up to two cells away from the bombardment center, suggesting CPC could move from one cell to another (Fig 2.5D). As we discussed, long-range repressors, *CPC* and *ETC1* (Kirik et al., 2004a), are directly activated by the TTG1-bHLH-MYB complex while the short-range repressor, *TRY,* is not. This may indicate that the accumulation of the active TTG1-bHLH-MYB complex in the trichome initials triggers primarily long-range inhibition but not short-range inhibition.

In addition, we also tested the movement potential of TTG1, GL1 and GL2 and we find that these proteins do not move in the leaf epidermis under the same condition where CPC moves. Another issue deserving special attention is that GL3 did not move from cell-to-cell in the leaf epidermis, in contrast to our earlier finding that GL3 moves

between root cell files (Bernhardt et al., 2005). By examining the protein accumulation and *Promoter::GUS* expression patterns, we find that GL3 is transcribed and the protein accumulates in the trichome initials and GL3 is transcribed in the surrounding epidermal cells but the GL3 protein is not detectable there (Fig 2.6). The apparent lack of GL3 movement in the leaf coupled with the transcription and protein pattern may indicate that the absence of GL3 protein in epidermal cells is not caused by GL3 trafficking into developing trichomes, but rather by some form of posttranscriptional or translational regulation.

The current model of trichome patterning is largely based on genetic analysis and molecular data obtained from the root hair system. The data presented in this paper demonstrate that a similar molecular mechanism by a TTG1-bHLH-MYB regulatory complex directly activating downstream targets is responsible for trichome patterning. Based on this mechanism, we have refined the model for trichome patterning. As shown in Fig 2.7, a functional activating complex TTG1-GL3/EGL3-GL1 activates trichome activators (*GL2* and *TTG2*) and single MYB repressors (*CPC* and *ETC1*) in the cell chosen to be a trichome. CPC and ETC1 then move into the neighboring cells where they, together with locally expressed repressors, compete with GL1 for binding to EGL3, forming an inactivating complex, TTG1-EGL3-CPC/ETC1. This inactivating complex disrupts the function of the activating complex. The decreased concentration of the TTG1-EGL3-GL1 complex in these surrounding epidermal cells is not enough to activate *GL2* and *TTG2* beyond a required initiating threshold level and the trichome cell fate is not triggered.

Our results also show differences between the trichome and root hair pathways at the molecular level: GL3 is preferentially transcribed in the cells where it functions during trichome development, while GL3 is transcribed in root hair cell files and accumulates and functions in non-root hair cell files during root hair patterning. This raises many new questions for this regulatory network. Identification of the molecular components which mediate the differentiation of bHLH expression patterns in different tissues will allow the study of how these key developmental complexes are regulated in the plant.

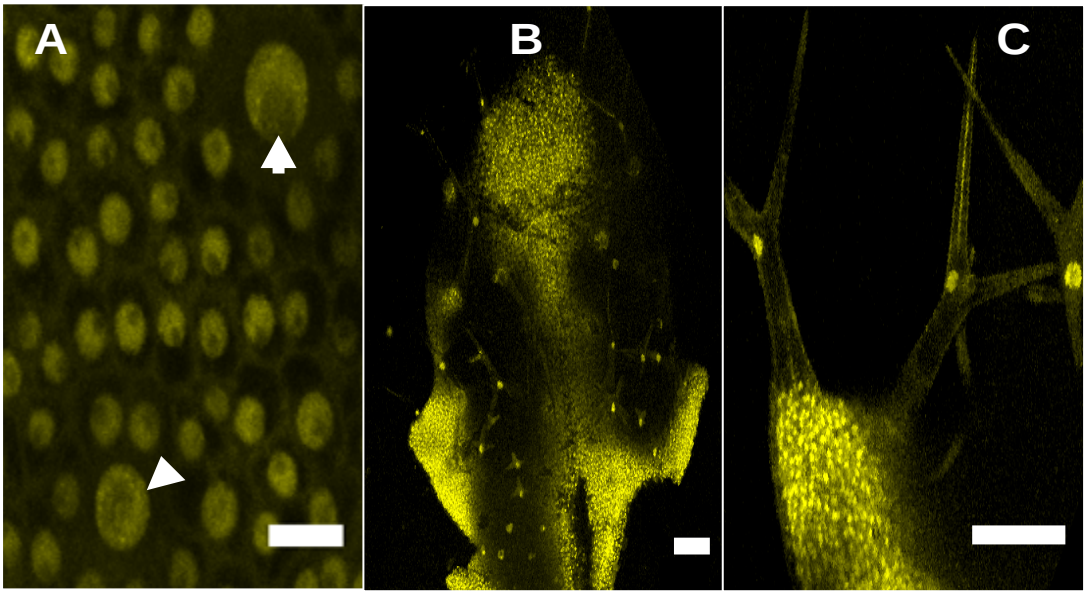

Fig 2.1. Expression pattern of YFP-TTG1 fusion in the leaf epidermis.

Maximum intensity projection images of confocal stacks of a TTG1::YFP-TTG1 construct in developing leaves of 20-day-old *ttg1* mutant seedlings. (**A**) Overview of a developing young leaf. This leaf is not flat so that in some areas the pavement cells are in focus and in other areas, focus is higher up on the trichomes. (**B**) Mature trichomes and surrounding cells. (**C**) Trichome initials (arrows) and surrounding cells. Bars in A = 10 μm; B = 100 μm; C = 50 μm.

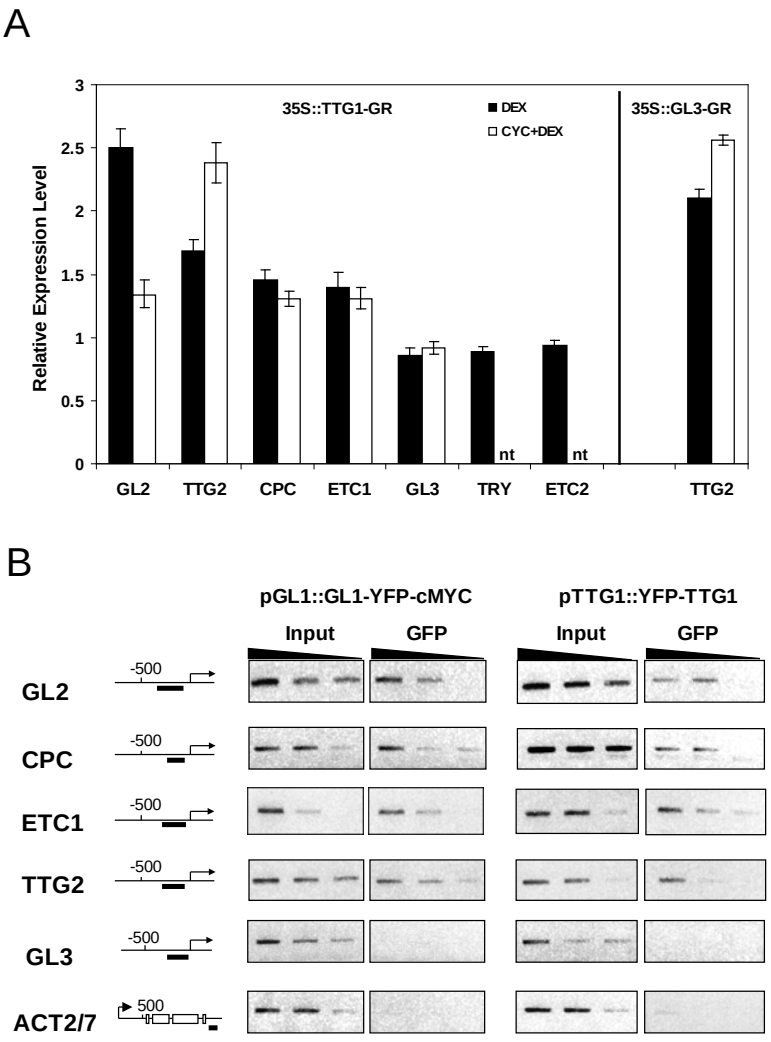

Fig 2.2. Direct activation of GL3 target genes by 35S::TTG1-GR.

(**A**) Gene expression levels were measured by relative quantitative PCR. The results were calculated using the comparative Ct method (ABI bulletin) and presented as fold changes compared to the mock or CHX treatment, which were standardized to the level of Actin expression. The induced expression levels of *GL2, TTG2, CPC* and *ETC1* were statistically significantly different from those of control treatments ($P < 0.05$); error bars indicate the ranges of expression change; nt: not tested. (**B**) Semi-quantitative PCR of ChIP experiments using *gl1*/pGL1::GL1-YFP-cMYC (left) or *ttg1*/pTTG1::YFP-TTG1 (right). PCRs were performed on three four-fold serial dilutions of the immunoprecipitated material, represented by the black slope.

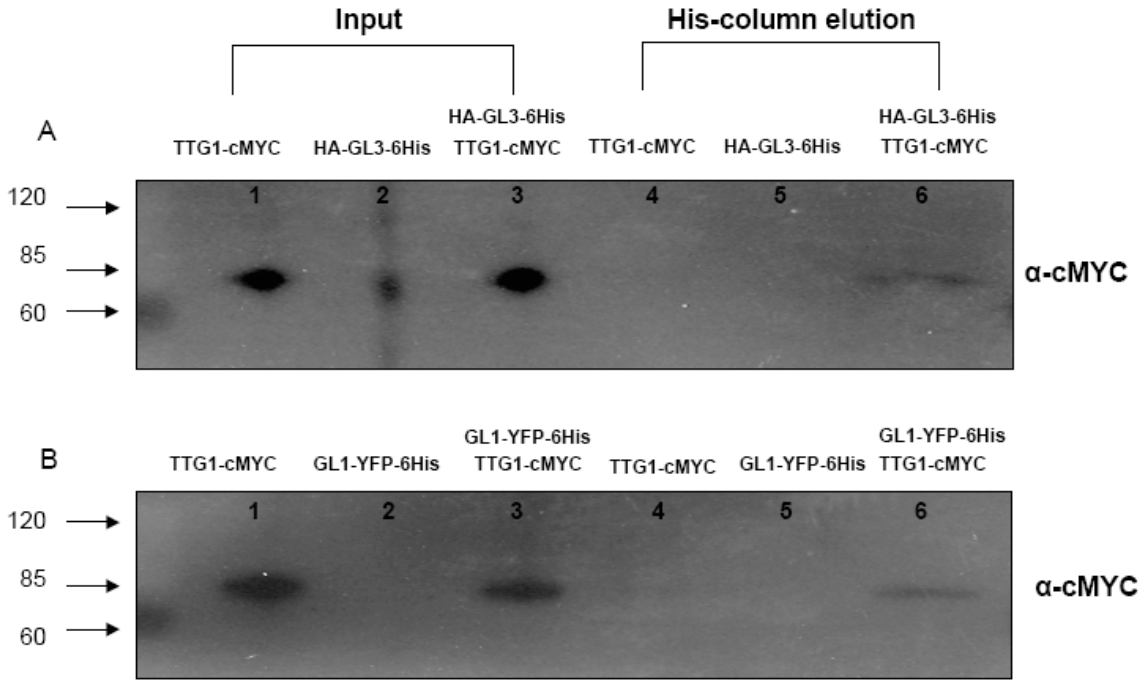

Fig 2.3. Co-precipitation of TTG1-cMYC with HA-GL3-6His and GL1-YFP-6His from seedling extracts.

His-Select Ni columns were used to pull down 6His-tagged fusion proteins. Input and eluted proteins were separated by SDS-PAGE gels in the order labeled. Membranes were probed with anti-cMYC mAb. In vivo interaction between TTG1 and GL3 is indicated in (**A**) and between TTG1 and GL1 in (**B**).

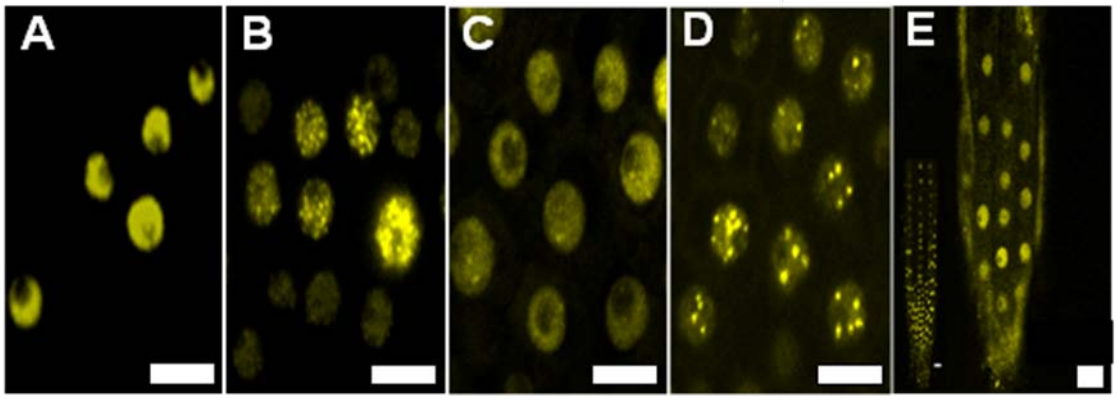

Fig 2.4. Speckled nuclear distribution of GL3 in the leaf epidermis of *ttg1* and *gl1* mutants.

Confocal images of (**A**) *ttg1*/pGL3::GL3-YFP, (**B**) *ttg1*/pGL1::GL1-YFP-cM, (**C**) *gl1*/pGL3::GL3-YFP and (**D**) *gl1*/pGL3::GL3-YFP. GL3-YFP is unevenly distributed and forms nuclear speckles in leaf epidermal cells of *ttg1* and *gl1* mutants (A and C). GL1-YFP formed only a couple of speckles in nuclei in occasional leaf epidermal cells (B). Uniform GL3-YFP distribution in root epidermal cells (inset picture shows GL3::GL3-YFP accumulation in hairless cell files forming wild type-looking stripes) (D). Bars = 10 μm.

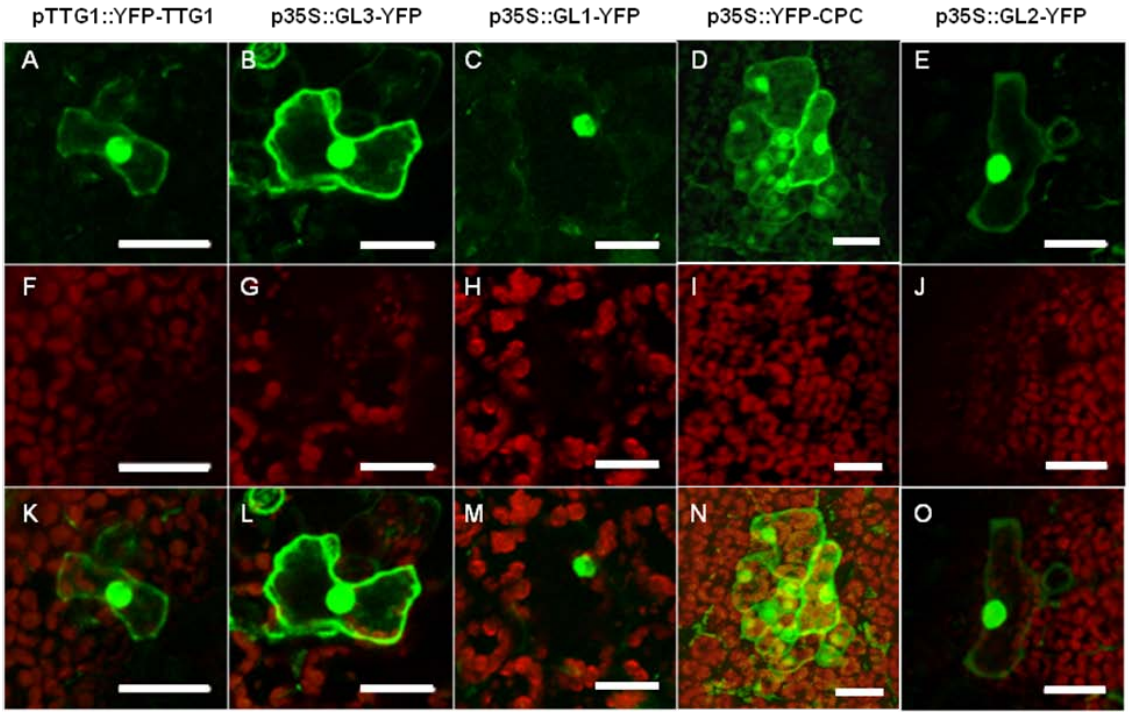

Fig 2.5. Intercellular trafficking of YFP-CPC in the leaf epidermis.

Confocal images of bombarded YFP fusion proteins. YFP is shown as green and chlorophyll autofluorescence as red. (**A-E**) YFP fluorescence. (**F-J**) Chlorophyll fluorescence. (**K-O**) Merged images. Only YFP-CPC shows cell-to-cell movement forming a cluster of fluorescent cells (D and N). TTG1, GL3, GL1 and GL2 fusions are cell autonomous, as fluorescence is restricted to single cells. Guard cells, as in the upper left of panel (B), are often autofluorescent under these conditions. Bars = 10 μm.

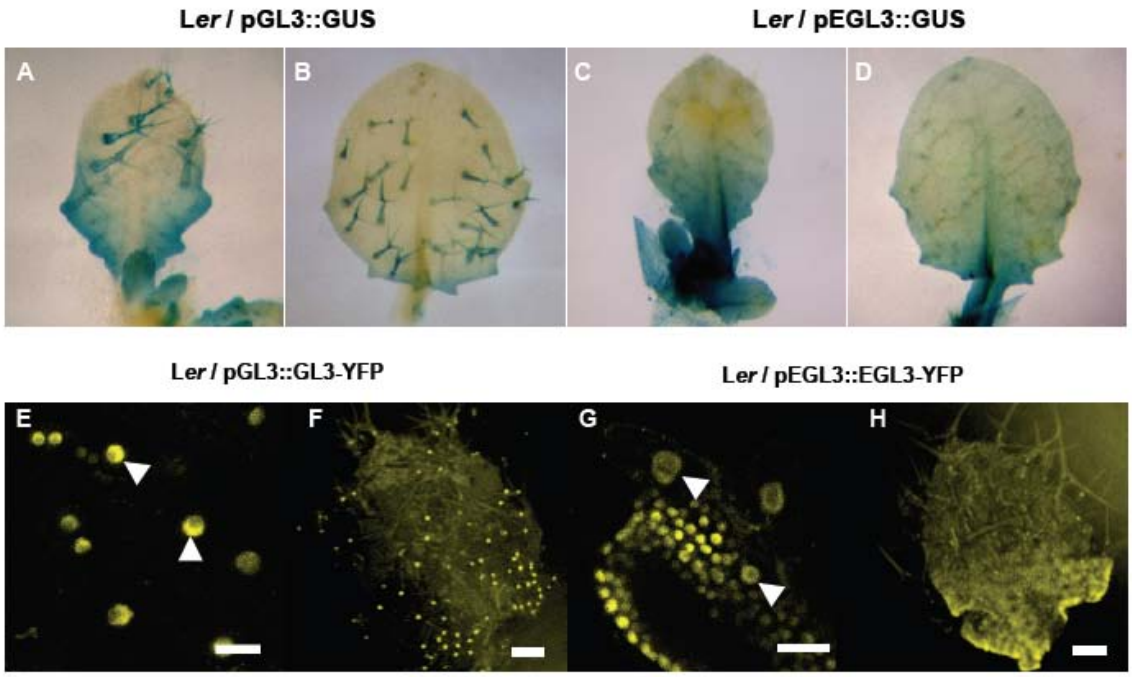

Fig 2.6. GL3 and EGL3 have overlapping but distinct expression patterns.

(**A-D**) Transcription patterns of *GL3* and *EGL3* promoter::GUS in wild type leaves. (**A**) and (**C**): Both *GL3* and *EGL3* are strongly transcribed in leaf primordia; high GL3 is observed close to the basal edge of the leaf while EGL3 is more wide spread and not restricted to the edge; trichome initials show higher levels of GL3 and EGL3 than surrounding epidermal cells. (**B)** and (**D**): strong GL3 expression becomes restricted to trichomes, while EGL3 expression persists in pavement cells as well as in trichomes. (**E-H**) Protein accumulation patterns of GL3 and EGL3 in wild type leaves. (**E)** and (**F**) GL3-YFP highly accumulates in trichome initials (arrows) and young trichomes. (**G**) and (**H**) strong EGL3-YFP was also detected in the trichome and non-trichome cells. Bar (E and G) = 20 μm; Bar (F and H) = 50 μm.

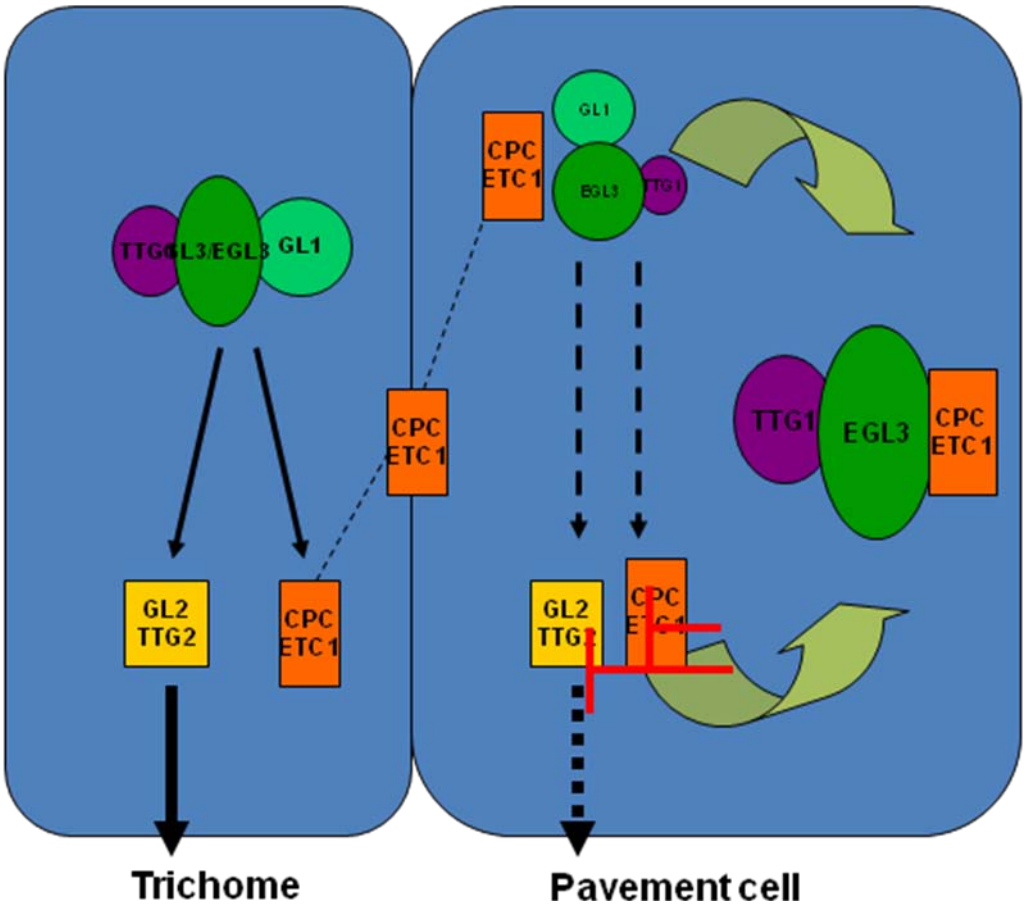

Fig 2.7. Model for *Arabidopsis* trichome/non-trichome cell fate specification.

Regulators of trichome fate are depicted in green shades, activators (GL2/TTG2) are in yellow and inhibitors (CPC/ETC1) are in orange. Black arrows indicate transcriptional activation. In trichome cells the inhibitors are directly activated by the activating complex and move (dashed lines) into neighboring cells, where they and endogenous inhibitors block the activity of the activating complex thereby decreasing the expression of GL2/TTG2 to below a required initiating threshold level (dashed arrows). Therefore, the trichome cell fate is not triggered.

Table 2.S1. Primers for plasmid construction

Clone	Primer name	Sequence
p5T3-5	BWD40X5	GGGAGCTCTTATCTTTTATTTATATTCTT TAGTTATGT
	BWD40X3	GGGAGCTCCTCGAGTAATTCAAACTGA CAAACTCCATTG
p5T3-5NX	RTTG1-N-INV	GCTCTAGAGGTGGTGGTATGGATAATT CAGCTCCAGATTC
	LTTG1-N-INV	GCTCTAGAGCGCGGCCGCAGGAAATGT GTGGTGCTGATTCG
pYFP-TTG1	YFP-TTG1-5'	AATGCGGCCGCTAATGGTGAGCAAGGG CGAGGAGC
	YFP-TTG1-3'	GCTCTAGACTTGTACAGCTCGTCCATGC CG
pTINV	RTTGINV	GGGGCGGCCGCCTCTAGAGGTGAGACTT TCTCTTTCGCTA
	LTTGINV	GGGGCGGCCGCCAACTCTAAGGAGCTGG ATTTTG
pTTG1-cMYC	TTGM5-1	AATGCGGCCGCGGAGTTATAGCTATGGA GCAAAAG
	TTGM3-1	GCTCTAGAGGCTCGAGAGGCCTTCAATT CAAG
pBGL1	GL1Pro5	GGGGTACCCCTCGAGTTAGTCAAAAATT AAAGTAG
	GL1Cod3	CCCAAGCTTGGCAAGGCAGTACTCAACA TCACCAG
pMter	GL1MYC5	CCCAAGCTTGGAGTTATAGCTATGGAGC AAAAGC
	GL1MYC3	CCTACGCGGAAGATATCGGCTCGAGAGG CCTTCAATTCAAG
	MYCGL1ter5	GAAGGCCTCTCGAGCCGATATCTTCCGC GTAGGTTTTCATT

	MYCGL1ter3	CGGGATCCCGGTATACCACTTTCATCTTA
pGL1-INV-cMYC	R-GL1-MYC-INV	AATATTTTACGTTTC CCCAAGCTTGGGCTCTAGAGCAGTTATA
	L-GL1-MYC-INV	GCTATGGAGCAAAAGC CCCAAGCTTGGCAAGGCAGTACTC
pGL1::GL1-YFP-cMYC	GL1-ENFP-5	CCCAAGCTTGGAGGAGGAGGAGGAATG GTGAGCAAGGGCGAGG
	GL1-ENFP-myc-3	GCTCTAGAGCCTCCTCCCTTGTACAGCTC GTCCATG
pGL1::GL1-YFP-6His	GL1-ENFP-5	CCCAAGCTTGGAGGAGGAGGAGGAATG GTGAGCAAGGGCGAGG
	GL1-ENFP-6His-3	GCTCTAGAGCCTAATGATGATGATGATG ATGTCCTCCTCCTCCTCCCTTGTACAGCT CGT
pGWHA-GL3-6His	35S-HA-GL3-6His-5-GW	GGGGACAAGTTTGTACAAAAAAGCAGG CTATGTATCCATATGACGTGCCGGACTA CGCCTCCCTCGGAGGAGGAATGGCTACC GGACAAAACAG
	35S-GL3-6His-3-GW	GGGGACCACTTTGTACAAGAAAGCTGG GTGCTGTCAATTTTAACTAAG
pGWGL3	35S-GL3-CFP-5-GW	GGGGACAAGTTTGTACAAAAAAGCAGG CTATGGCTACCGGACAAAACAG
	35S-GL3-CFP-3-GW	GGGGACCACTTTGTACAAGAAAGCTGG GTAACAGATCCATGCAACCCTTTG
pGWGL1	35S-GL1-CFP-5-GW	GGGGACAAGTTTGTACAAAAAAGCAGG CTATGAGAATAAGGAGAAGAG
	35S-GL1-CFP-3-GW	GGGGACCACTTTGTACAAGAAAGCTGG GTAAAGGCAGTACTCAACATCAC
pGWCPC	35S-YFP-CPC-5-GW	GGGGACAAGTTTGTACAAAAAAGCAGG CTTCATGTTTCGTTCAGACAAGG
	35S-YFP-CPC-3-GW	GGGGACCACTTTGTACAAGAAAGCTGG GTTCATTTCCTAAAAAAGTCTC

pGWGL2	35S-GL2-YFP-5-GW	GGGGACAAGTTTGTACAAAAAAGCAGG CTTGATGTCAATGGCCGTCGACATG
	35S-GL2-YFP-3-GW	GGGGACCACTTTGTACAAGAAAGCTGG GTGGCAATCTTCGATTTGTAGACTTC
pLBJ131-2	EGL3-YFP-5	GGGGACAAGTTTGTACAAAAAAGCAGG CTTTGTTTATGACCATACACGTGG
	EGL3-YFP-3	GGGGACCAGTTTGTACAAGAAAGCTGG GTCACATATCCATGCAACCCTTTG

Table 2.S2. Primer pairs for quantitative PCR

Gene	Primer name	Sequence
GLABRA2 (GL2)	GL2-RT-F	GGACAGCAACACGGAGAAGGAG
	GL2-RT-R	TGTCCTCGATGATGCAACCG
TRANSPARENT TESTA GLABRA2 (TTG2)	TTG2-RT-F	ATTACAAATGCACACACCCG
	TTG2-RT-R	TCTGAAACTTGACCTTCCACTGA
CAPRICE (CPC)	CPC-RT-F	CTGCTTGCCGAATATCATGGTG
		AGCCAAGGCTTCTTGTTCCG
	CPC-RT-R	TGACATCTTCACAGCTTCCCATT
ENHANCER OF TRY and CPC1 (ETC1)	ETC1-RT-F	AAGACCAATCCAACCATTGTTG
	ETC1-RT-R	TCTTCCCACTCAAGACTGCTCA
GLABRA3 (GL3)	GL3-RT-F	GATCAGCTTGGTCTACGGAGGAG
	GL3-RT-R	CAGCGACGGAGAGAGACTCG
TRIPTYCHON (TRY)	TRY-RT-F	AGGATAGATAGAAAAGCGAGGACG
		CTGCTTGCCGAATATCATGGTGA
	TRY-RT-R	TAAACGACGCCAGCTTCACTC
ENHANCER OF TRY and CPC2 (ETC2)	ETC2-RT-F	CAAACTAAGTTCACTCGATCCCG
	ETC2-RT-R	AAACTCCCATTCGATGCTACTCA
ACTIN (ACT)	ACT-RT-F	TCCATTCTTGCTTCCCTCAG
	ACT-RT-R	ATCATACTCGGCCTTGGAGA

Table 2.S3. Primers for ChIP

Gene	Primer name	Sequence
GL3	GL3pro-A1	AAACGGCAACTGTTTCATCA
	GL3pro-B1	TTCTGTTTTGTCCGGTAGCC
GL2	GL2-A1	GTAGCTGAAATTGGAAGCTGATA
	GL2-B1	ACTGCTCTTGACTTTTAGGTGCTT
CPC	CPC-CHP-5	CAAGAACACATAGAAGGGAC
	CPC-CHP-3	CATAGAGAAAGAAGAACGAC
ETC1	ETC1-A1	GAGCATTGCACACATACTGACATA
	ETC1-B1	TATCAATCAATACGGTTTGGTACG
TTG2	TTG2-A1	ATCACAACTTCAGTTTTTGCATTC
	TTG2-B1	TTTGCTATTACCACTCTTTCACCA
ATC2/7	Act-JP1595	CGTTTCGCTTTCCTTAGTGTTAGCT
	Act-JP1596	AGCGAACGGATCTAGAGACTCACCTT

Chapter 3: Examining *GLABRA2*'s Function

Glabra2 (*GL2*) has long been identified as an important regulator of epidermal cell fate and differentiation, yet much of how the GL2 protein functions remains unknown. GL2 has two identified domains, a homeodomain of the ZLZ type (leucine zipper loop leucine zipper) and START lipid binding/transfer domain, which characterize both it and its family. The homeodomain presumably facilitates GL2's role as a transcription factor, while the START domain has no known function in GL2 or any other plant protein. These regions only encompass two thirds of GL2 leaving a large C-terminus with no identified domains, yet in a related family of HD-STARTs, a protein-protein interaction domain has been identified (Chandler et al. 2007). The C-terminus of GL2 is highly conserved across its family and even across species. Additionally analysis of GL2's sequence shows that there is an acidic region likely to be involved in its role as a transcription factor. In this chapter I will discuss work done to identify how GL2 functions through fragment analysis, protein-protein interaction studies, and ectopic expression of GL2 and other trichome regulatory genes.

RESULTS

Overexpression of GL2

GL2 mutants have disrupted trichome, root hair, and seed coat development. However when GL2 is overexpressed using CaMV35S promoter in wild type plants, a dominant negative phenotype is produced. The overexpression of GL2 disrupts trichome development in a manner similar to *gl2* mutant. It also disrupts root hair, mucilage, and

seed coat development. The same phenotype occurs with N- or C-terminal tags. Overexpression in *ttg1-1, gl1-1, gl3-1, gl2-1, gl3/egl3/tt8/myc1, try, cpc*, and *try/cpc* had either no trichome effect on lines without trichomes or a dominant negative trichome phenotype on ones with trihcomes. Overexpression of GL2 in *ttg2-1* mutant resulted in a seed where the embryo ruptures the seed coat prematurely. These seed coats appear to be more fragile than normal and do not appear able to grow to a size sufficient to cover the embryo (Fig 3.1). While these seeds have ruptured seed coats and even some embryos completely falling out of the seed coat, the embryos are still viable indicating that normal embryo maturation and dehydration is not comprimised. Trichomes on these plants have either no new phenotype, or a slight dominant negative phenotype. Double mutants of *gl2-pi* and *ttg2-1* have an increase in abnormal seed shape from either parential line, and a number of embryos spontaneously coming out of the seed coat in a manner reminiscent of GL2 overexpressed in TTG2 which is generally not seen in either single mutant.

Transformation of *gl2-1* and Col plants with GL2 under the control of its own promoter does not produce the dominant negative phenotype. This construct complements *gl2-1*. As described previously (Ohashi et al. 2002) there seems to be increase in the formation of clusters of adjacent trichomes in these transgenic lines.

Overexpression of TTG1 complex members in *gl2-1*

GL2 mutants have a unique phenotype when compared to other trichome mutants. No other mutants have been identified that halt development in the elongation stages. There are many other trichome genes that when overexpressed will increase trichome

elongation and branching and I hypothesized that there could be suppression of the *gl2* mutant phenotype by overexpressing some of these genes in *GL2* mutant lines.

GL3 and EGL3 overexpressed in *gl2-1* produced an increased number of trichomes, as they do in wild type plants. Both also increased expansion or outgrowth of the nub-like trichomes of *gl2-1* slightly and had the effect of initiating or increasing branching on the *gl2-1* trichomes that were more advanced with odd patterns and shapes occasionally emerging.

Overexpression of GL1 in *gl2-1* resulted in suppression of the *gl2* mutant phenotype producing trichomes that expand and branch on the first true leaves. There is also an increased and irregular branch pattern (Fig 3.2).

Overexpression of TT8, MYB5, MYC1, TTG2 and MYB106 had no noticeable effect on *gl2-1* mutant plants.

Analysis of GL2 protein

GL2 consists of a homeodomain, acidic region, START domain and conserved C-terminus; however the role of each domain is largely unknown. A series of deletions were created in attempt to assign a function or role to the other domains/regions (Fig 3.3). Regardless of which GL2 protein fragment was overexpressed, a dominant negative phenotype was displayed. This held true with fragments that contained C- or N-terminal tags, as well as untagged versions. The severity of the phenotype was slightly decreased in constructs missing the homeodomain, however it was still present. The phenotype encompassed all three epidermal characters that GL2 controls: trichomes, the outer seed coat, and root hairs (Fig 3.4 and 3.4).

GL2 nuclear Localization

GL2 as a transcription factor is normally localized in the nucleus. Analysis of its sequence shows a possible nuclear localization signal RKRKK at amino acid position 99. The deletion constructs from the previous section allowed analysis of which part of GL2 is required for its nuclear localization. In constructs that contain the predicted nuclear localization signal (NLS) GL2 is localized in the nucleus, while the constructs missing the NLS are localized throughout the cell (Fig 3.6).

GL2 antisense

GL2 mutants have been described in several papers as possibly not being null mutants. To examine this possibility in a new way, an antisense version of GL2 was created in Col wild type. These transgenic lines looked like *gl2-5* (*GL2* mutant in col background) or like a slightly less severe *gl2-1* mutant. This shows in another way that the known mutants of *gl2* are at least reduction of function and not gain of function or neomorphic mutations.

Dimerization of GL2

GL2 is predicted to homo or heterodimerize through its homeodomain, however this has never been demonstrated. In yeast two hybrid experiments, full length GL2 does not interact with itself (Fig 3.7). However when the START domain is removed from GL2 (GL2ΔS), GL2 is able to dimerize in yeast. The interaction is greatest between two GL2ΔS constructs, however some interaction is present between a full length GL2 protein and a GL2ΔS protein (Fig 3.7). The C-Terminus fragment which lacks the homeodomain was not found to bind either to itself or to full length GL2.

GL2 Binding Partner Screen

An attempt at finding GL2 binding partners was made, using the GL2ΔS construct that has been shown to be functional in yeast, as the bait for a yeast two hybrid screen of a cDNA library. In a screen of approximately 580,000 colonies 19 initial positives were identified. Of those 19, only 1 (Fig 3.8) made it through the series of tests verifying the interaction (including check of autoactivation and vector swap). This possible GL2 binding partner is At5g17350 an unknown protein which has one close (At3g03280) and another moderate homolog (At4g02090). At3g03280 was tested in yeast two hybrid experiments and found to not interact with GL2. Analysis of T-DNA lines showed no phenotype for insertions in any of the three genes. Overexpression of At5g17350 and At3g03280 in wild type and gl2 mutant also displayed no obvious altered phenotype.

MATERIALS AND METHODS

Overexpression Analysis

GL2 full length, fragments, and antisense constructs were created using GL2 cDNA from Columbia (Col) ecotype with Invitrogen's gateway recombination cloning system. Primers for each construct are listed in Table 3.1. Untagged versions were cloned into pB7WG2, N-Terminal tagged in pB7WGY2, and the C-terminal tagged in pB7YWG2 (Karimi et al. 2002). The antisense vector, LBJ17 RFB REV, was created by cloning a Gateway destination vector cassette in reverse orientation in relation to the 35S promoter of LBJ17 (Payne et al. 2000). The GL2 full length coding sequence was recombined into LBJ17 RFB REV to create the 35S-antisense GL2 construct.

GL2ΔS construct was created using Site-directed mutagenesis kit from Stratagene using a full length GL2 construct in pDONR Zeo (pDZ GL2) with primers available in Table 3.1. This created a gateway DONR vector for later recombination reactions into destination vectors.

GL3, EGL3 and MYB5 constructs were previously decribed (Zhang et al. 2003; Gonzalez et al. 2008). GL1, TT8, MYC1, TTG2 and MYB106 constructs are full length from start to stop cloned into pB7WG2.

Yeast Two Hybrid

All yeast two hybrid constructs were created with Gateway using cDNA from seedlings of the Columbia ecotype. The binding domain vector utilized was pGBT9 RFB which was created by inserting a gateway destination vector cassette into pGBT9. The activation domain vector used was pACTGW (Nakayama et al. 2002). The screen was performed using the GL2ΔS fragment in pGBT9 RFB as the bait, against the Kim and Theologis λ-ACT 2-hybrid library available from ABRC (Kim et al. 1997). Primers for the possible binding partner for GL2 are listed in Table 3.2.

DISCUSSION

The analysis of the overexpression of GL2 fragments indicates that all parts of the GL2 protein can disrupt proper trichome development in a dominant negative manner. It is not surprising that the fragments containing the homeodomain can interfere with normal development as this region contains both a DNA binding domain and protein dimerization domain. However fragments lacking both the homeodomain and the nuclear localization signal show that the dominant negative phenotype is not dependent on being

in the nucleus, binding DNA, or dimerizing through the homeodomain. The dominant negative phenotype from fragments containing just the START domain and/ or the C-terminus show the likelihood that these regions of the protein have important interaction roles in GL2 function. Since the C-terminus of the protein is well conserved it is likely that it has an important function in protein-protein interaction. In the closely related HD-ZIP family of homeodomain proteins which also contain START domains, there is a protein binding domain identified in the C-terminus of the protein (Chandler et al. 2007). Interaction with a protein through the C-terminus could cause some of the dominant negative phenotypes. The START domain might also have a role through its lipid binding ability. The ligand could be soaked up by the overexpressed protein preventing the ligand from performing its normal role in GL2 function, whatever that may be.

While there still might be some sort of RNAi effect responsible for the dominant negative phenotype, at least some protein is seen in these lines as evidenced by visualization of the YFP tag. If RNAi type suppression is the case it would invoke the argument that GL2 is more RNAi sensitive than other regulators or more specifically targeted by RNAi than any other trichome regulatory gene known. This seems highly unlikely.Through these experiments several predicted characteristics of GL2 were finally experimentally verified. GL2 is known to be nuclear localized (Szymanki et al. 1998) yet the signal for this localization has not been identified *in vivo* until now. The same goes for the homodimerization of GL2. It has been suggested for a long time yet not proven. What makes these results surprising is the inability of full length versions of GL2 to dimerize while GL2 lacking the START domain strongly dimerizes. This suggests the

START domain may have a role in regulating dimerization or its ability to activate transcription in response to the presence or absence of an unidentified ligand. If this conclusion is correct, a screen for possible ligands for GL2's START domain might be possible using the yeast two hybrid experimental system described. In this case, yeast would be treated with potential ligands to attempt to allow dimerization of full length GL2 containing the START domain, activating the yeast reporter genes. Presently, it is unclear whether the negative yeast result is due to the lack of interaction or the strong repression of transcription. GL2 has been reported to both activate and repress different targets in root hair development (Ohashi et al. 2003; Tominaga-Wada et al. 2009), so either scenario is possible. The START domain might facilitate the switch between repression and activation of its targets, in which case the ligand screen could work here too. However, the switch between activation or repression of GL2's transcriptional targets could also be accomplished through interaction with different binding partners.

While the screen for possible binding partners of GL2 did not lead to discovery of any particularly interesting genes, it did find a single possible binding partner. This gene may have a role in GL2 function that has not yet been discovered. The screen was also in no way comprehensive. One outstanding problem is the unavailability of an optimal library. The library used in this screen is made from cDNA from 3 day old etiolated seedlings. While such seedlings do have GL2 expressed, there are no trichomes present in the tissue so that finding binding partners that affect GL2's role in trichome development is less likely.

An interesting result of overexpression of different trichome genes in *gl2-1* mutant plants was the finding that the early nub trichome phenotype was unable to be

suppressed except by the MYB, GL1. This is made all the more interesting in that GL1, when overexpressed in most backgrounds, results in a dominant negative phenotype (Schnittger et al. 1998). It is the only major trichome gene, other than GL2, known with this phenomenon. The only other reported way to increase elongation of these nub trichomes is in *gl2 try* double mutants (Hulskamp et al. 1994). This is consistent with the GL1 overexpression experiment as the MYB, TRY, is a known antagonist of GL1 (Schnittger et al. 1998).

TTG2 is known as a regulator of seed size in addition to regulating mucilage production and seed coat development. It is interesting that overexpression of GL2 in the *ttg2-1* background produces what appears to be either smaller or less durable seed coats. Due to the dominant negative results in all other GL2 overexpression experiments, it is assumed that GL2 is acting as a dominant negative in this case also. Further analysis of *ttg2 gl2* double mutants is required to verify this conclusion but the *gl2-pi ttg2-1* mutants appear to also have abnormal seeds. This double mutant combination is complicated by the fact that they are in different ecotypes, which might account for some of the variation in the seeds. It is not clear what is going on within the seed to make the seed coat tear and rip in the GL2 overexpressed in *ttg2-1* line. Both genes are important to seed coat development and both are expressed in the embryo. Some work performed by Antonio Gonzalez of Alan's Lloyd lab has found evidence for abnormal packaging of the embryo in *ttg2* and other seed coat mutants (unpublished results). So the tearing of the seed could be the result of abnormal embryo size, abnormal embryo folding, or from a modification of the seed coat.

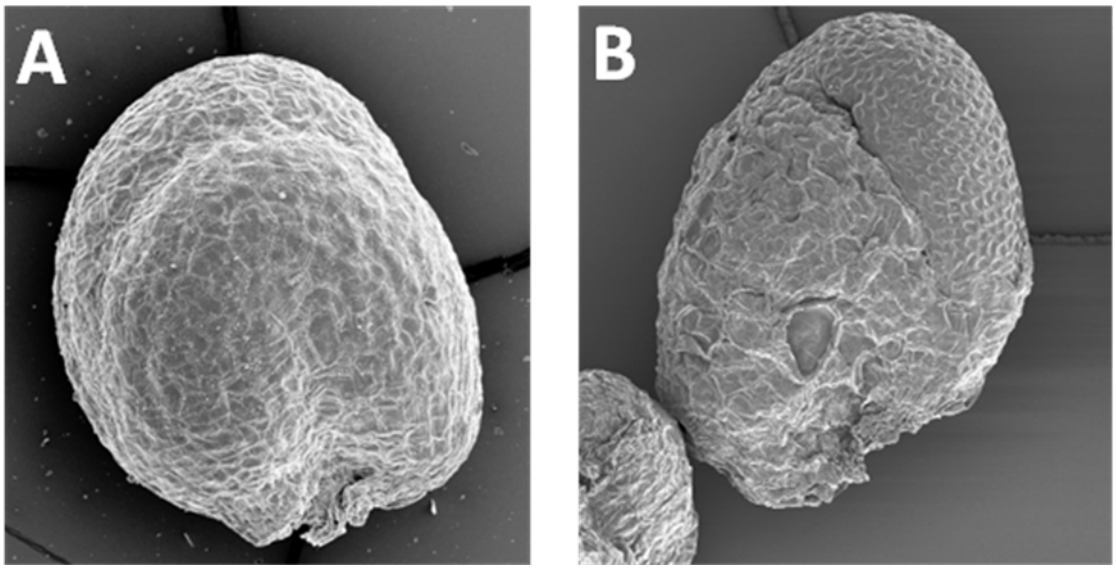

Fig 3.1. Overexpression of Gl2 in *ttg2-1*.

SEM of seeds from A.) *ttg2-1* untransformed, B.) 35S:GL2 in *ttg2-1* with ruptured seed coat.

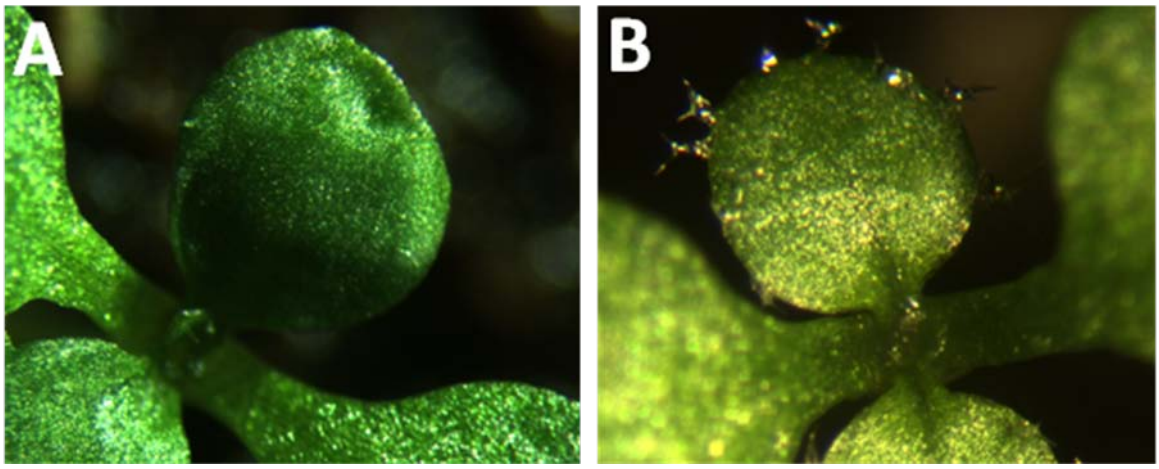

Fig 3.2. Overexpression of Gl1 in *gl2-1*.

First true leaves of A.) *gl2-1* untransformed. B.) 35S:GL1 in *gl2-1*.

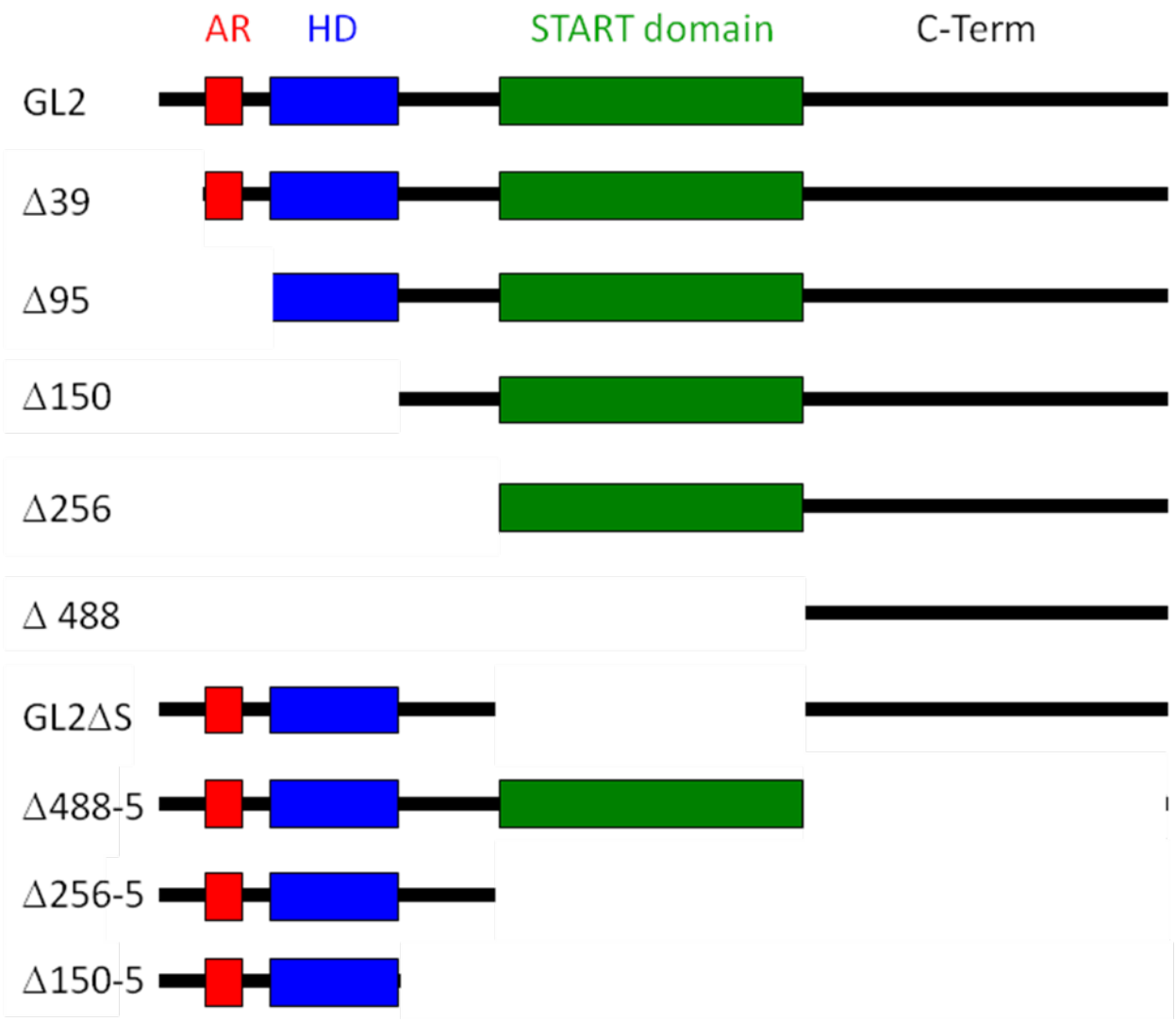

Fig 3.3. Fragment analysis of GL2.

Showing the four areas of interest within GL2 protein: the acidic region (AR) in red, homeodomain (HD) in blue, START domain in green, and the conserved C-Terminus.

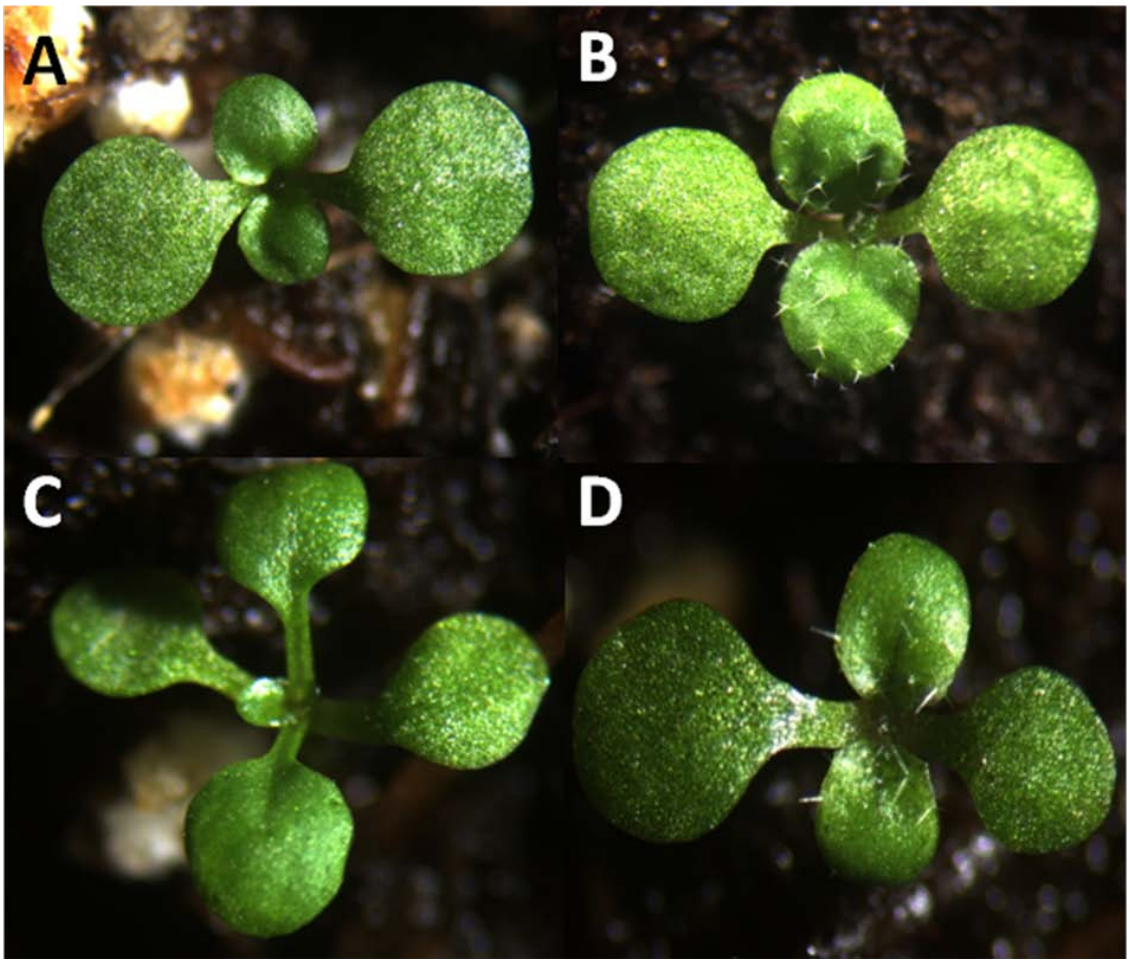

Fig 3.4. Overexpression of GL2 Fragments Trichome Phenotypes.

Overexpression of GL2 fragments results in dominant negative trichome phenotype. Dissecting scope images of A.) *gl2-1*, B.) Col wild type, C.) Col overexpressing D95, and D.) Col overexpressing D488.

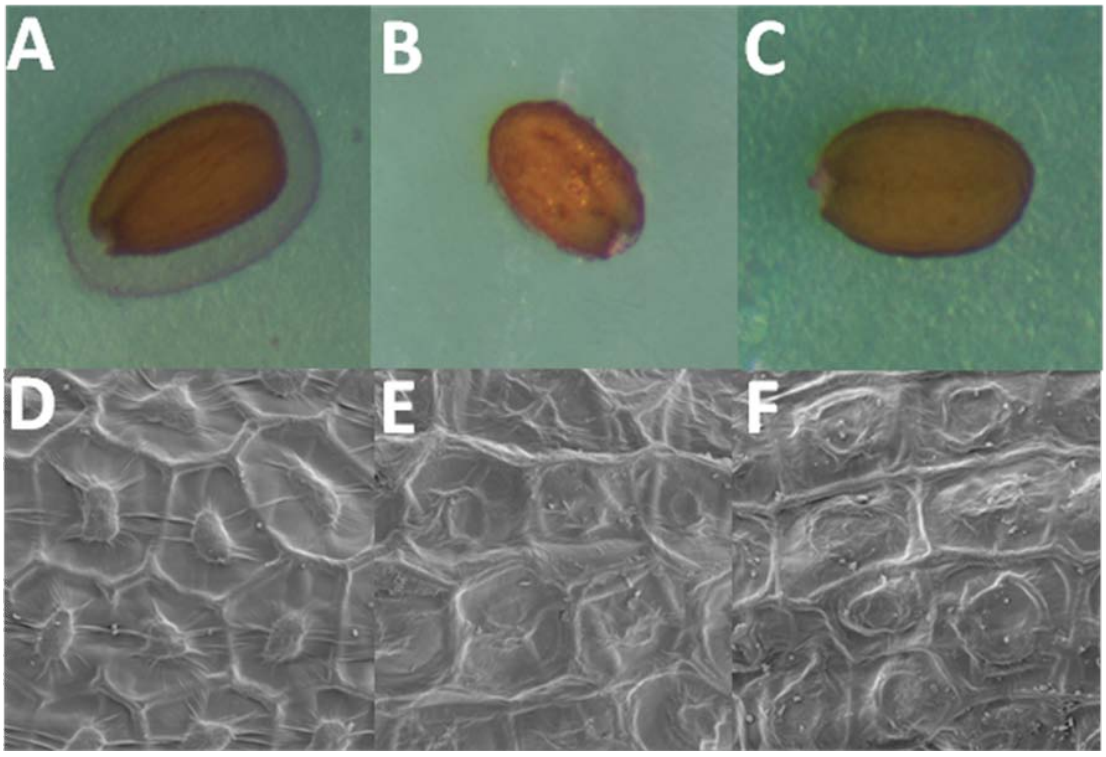

Fig 3.5. Overexpression of GL2 Fragments Seed Coat Phenotypes.

Overexpression of GL2 fragments produces a dominant negative phenotype in seed coat development. Ruthedium Red stained mucilage in A.) wild type Col, B.) Col overexpressing D95 and C.) *gl2-1*. SEM of D.) wild type col, E.) col overexpressing D95 and F.) *gl2-1* seed coats.

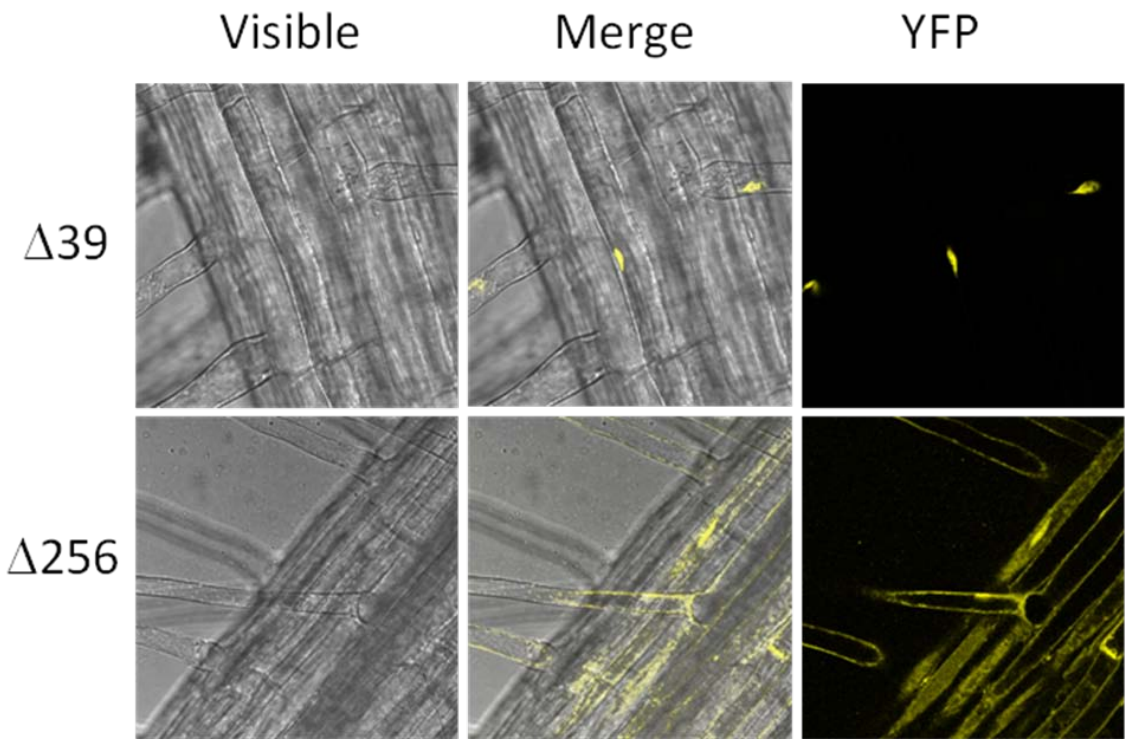

Fig 3.6. GL2 protein localization.

GL2 fragments with the NLS at amino acid position 99 are nuclear localized as in the D39 construct. Without the NLS as in D256 GL2 is localized throughout the cell.

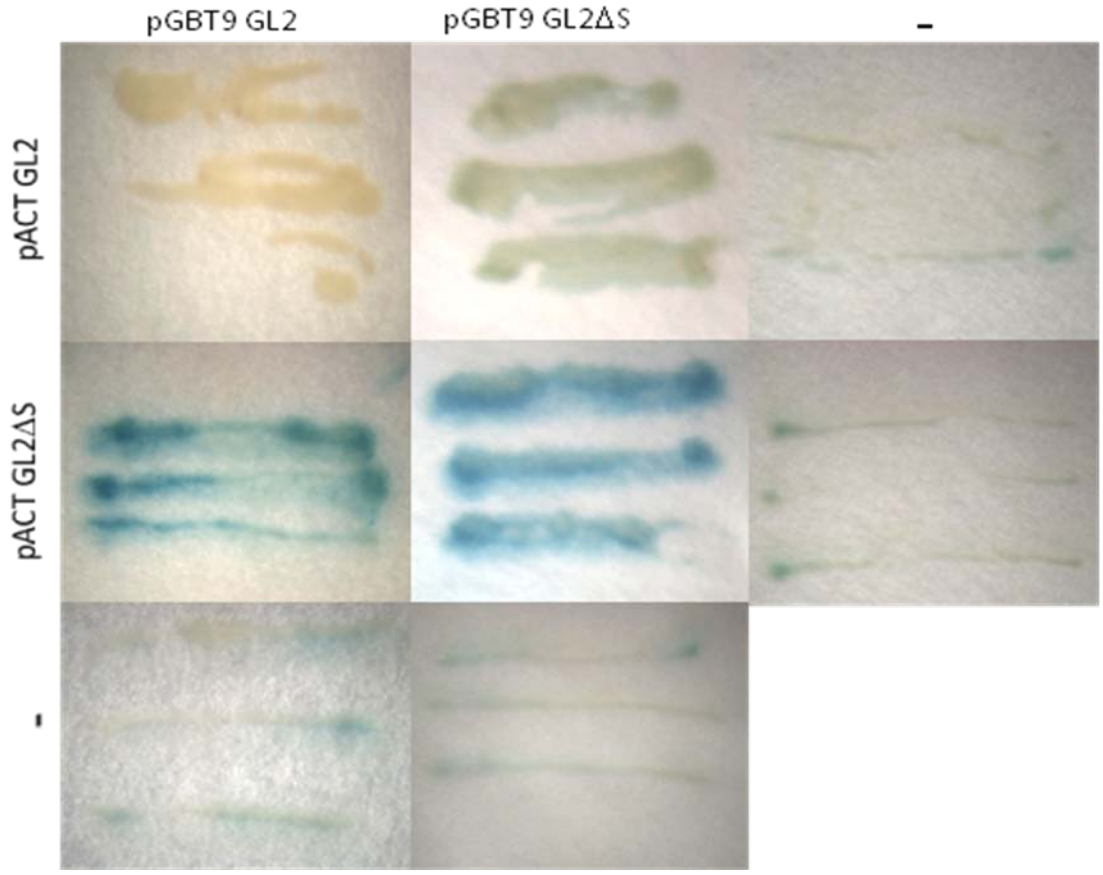

Fig 3.7. Test of GL2 dimerization.

Yeast Two Hybrid analysis showing Gal staining of either full length GL2 or GL2 without its START domain (GL2DS) in activation domain (pACT) and binding domain (pGBT9).

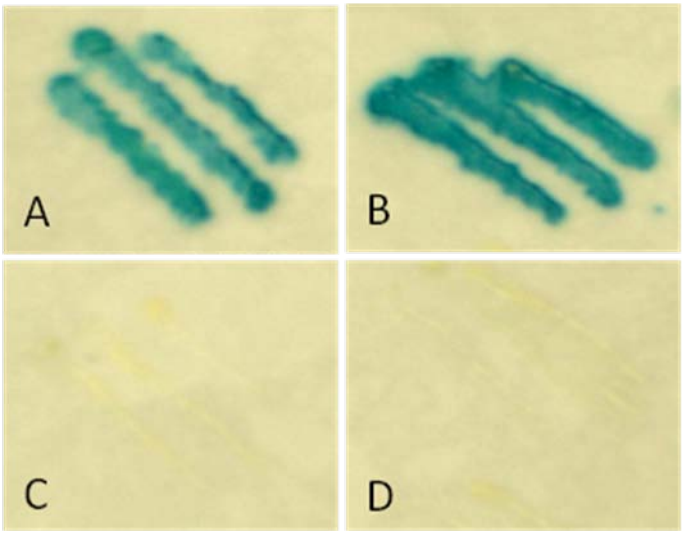

Fig 3.8. Yeast Two Hybrid Analysis of At5g17350 interaction with Gl2.

A.) pACT GL2DS/pGBT9 At5g17350, B.) pACT At5g17350/pGBT9 GL2DS, C.) pGBT9 At5g17350 and D.) pACT At5g17350.

Table 3.1 Primers for Cloning GL2.

Primer Name	Sequence
GL2-Start	ggggacaagtttgtacaaaaaagcaggcttgATGTCAATGGCCGTCGACATG
GL2 Stop	ggggaccactttgtacaagaaagctgggtgTCAGCAATCTTCGATTTGTAGAC
GL2-P2	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAATCATCTGCTACAGTTATACC
GL2-P3	Ggggacaagtttgtacaaaaaagcaggcttgacgtacgtattatacggacgg
GL23NS	GGGGACCACTTTGTACAAGAAAGCTGGGTGGCAATCTTCGATTTGTAGACTTC
GL2-39	ggggacaagtttgtacaaaaaagcaggcttgATGAACCCTGAGGAGGATTTCCTG
GL2-95	ggggacaagtttgtacaaaaaagcaggcttgATGAAGGGCACTAATAAGAGAAAGAG
GL2-160	ggggacaagtttgtacaaaaaagcaggcttgATGCAAGAACGGCACGAGA

	ACTC
GL2-245	ggggacaagtttgtacaaaaaagcaggcttgATGTTCTACACGGGCGTCTTTGC
GL2-488	ggggacaagtttgtacaaaaaagcaggcttgATGGTCTTCTTCATGGCTACCAAC
GL2-488-5	ggggaccactttgtacaagaaagctgggtgATGAAGGCGTTCGCAATGGAGCTG
GL2-245-5	ggggaccactttgtacaagaaagctgggtgGCAAAGACGCCCGTGTAGAA
GL2-160-5	ggggaccactttgtacaagaaagctgggtgATGATAGCCTTGATCTGTGTGCGG
GL2-LS-SDM 1	CACGGGCGTCTTTGCCCTCGAGATCTTCTTCATGGCTACCAACGTCCCC
GL2-LS-SDM 2	GGGGACGTTGGTAGCCATGAAGAAGATCTCGAGGGCAAAGACGCCCGTG

Table 3.2 Primers for possible GL2 binding partners.

Primer Name	Sequence
AT5G17350 5' fwd	ggggacaagtttgtacaaaaaagcaggcttgATGGGAAACTATGTTTCTTCTGC
AT5G17350 3' rev	ggggaccactttgtacaagaaagctgggtgTTAAGACAAAGATTCTTCAACTATAGTTTC
AT3G03280 5' fwd	ggggacaagtttgtacaaaaaagcaggcttgATGGGCAACTACGTTTCATG
AT3G03280 3' rev	ggggaccactttgtacaagaaagctgggtgTTAGGGCACATGATCTTCAGC

Chapter 4: Identification of GLABRA2 and TRANSPARENT TESTA GLABRA2 Transcriptional Targets

The TTG1 combinatorial transcription factor complex is required for all stages of trichome development. However, its main transcriptional targets are themselves transcription factor loci, *Glabra2 (GL2)* and *Transparent Testa Glabra2 (TTG2)*. GL2 and TTG2 likely target the structural genes required to take the trichome cell through its developmental plan. Since there are currently no identified transcriptional targets for either GL2 or TTG2 in the trichome pathway, there is no direct link from the TTG1 complex through GL2 and TTG2 to the actual emergent trichome phenotype. Identifying targets for GL2 and TTG2 should provide that link and supply insight into what it takes to differentiate from a pavement cell to a trichome cell. Using bioinformatics I have isolated a set of possible targets of GL2 and TTG2, and many of the potential GL2 targets have been experimentally tested.

RESULTS

Identification of GL2 targets using correlation data

To identify possible targets of GL2, candidates were isolated from publicly available expression databases through the use of expression correlation studies. I used Expression Angler from Bio-Array Resource (BAR) to calculate the Pearson correlation coefficient for all genes in a given data set compared to GL2. Pearson correlation coefficient measures how similarly two genes respond across a series of expression data samples. It does not refer to the magnitude of the response but rather the direction of the

response compared to the other data points. The Pearson correlation coefficient is expressed as a number from 1 to -1 with 1 being perfect correlation, zero being no correlation and -1 being perfect negative correlation. Using Expression Angler, I constructed a set of 300 genes from several expression data sets (AtGenExpress Consortium's Tissue series, Botany Array Resource, and NASCArrays) that had high positive correlation to GL2. The most correlated gene was At3g24250 a glycine-rich protein with a correlation coefficient of 0.969 in the NASCArrays data set. I also collected a set of anticorrelated genes, however they seemed less promising and were not examined further. The top anticorrelated gene was At3g50830 or ATCOR413-PM2 (Cold-Regulated 413-Plasma Membrane 2) with a correlation coefficient of -0.554 in the Botany Array Resource data set. Three TTG1 complex members, MYB23, MYB5, and TT8 were identified in the 300 positively correlated genes.

Expression correlation with GL2 is not sufficient evidence that a gene is an actual target of GL2. Corellation could mean that an upstream gene or a gene parallel to GL2 regulates the potential target, thus further analysis of the correlated genes is required. The GL2 family of homeodomain/START domain proteins has an identified binding element, L1-box, which is also present in the known targets of GL2 in the root hair pathway. Analysis of the promoters of the correlated genes for the presence of an L1-box would give a stronger suggestion for GL2 binding and therefore control of the gene's transcription. To accomplish this I used Promomer from BAR which will search 1000 bp upstream of a set of genes for a given DNA sequence.

The genes that were highly correlated and had an L1-box are all candidates to be targets of GL2. However as the correlation is across the whole data set and includes

tissue with no trichomes such as roots, seeds and flowers this list is not necessarily a list of possible trichome targets of GL2. To verify that the candidate is a realistic trichome target I examined the expression data visually using AtGenExpress' Expression Visualization Tool. The candidates' expression was qualitatively analyzed for seemingly significant expression in leaf tissue. Also, as there are expression data arrays from *gl1* mutant plants in the expression sets, I looked for a decrease in expression in the gl1 data points compared to similar wild type samples. My reasoning was that GL1 positively regulates GL2, therefore GL2 regulated genes should be down regulated in *gl1* mutants. There were several genes that had unacceptable leaf expression levels, that had data suggesting root or seed targeting by GL2. There were also genes with expression levels that were at very low levels across all data points and deemed not to be expressed enough to be the type of target desired in this screen. After these three screens, 13 genes including MYB23, MYB5, and TT8 were identified as possible targets of GL2 (Table 4.1).

Identifying Possible Targets using Published trichome specific Microarry data

An additional source of potential targets of GL2 was a set of trichome cell specific microarray results from Jakoby et al. (2008). This data includes expression data for isolated mature trichome cells, pavement cells, *gl3* mutants, and *try* mutants. Mining this data began with sorting for genes that were upregulated in trichomes at least 2 fold. Next, because GL2 is generally restricted to trichome cells in mature leaves, its targets should be lowly expressed in the rest of the leaf. The data was sorted for leaf expression at less than 15 arbitrary expression units on a linear scale based on the results of the

microarray (it was determined that a gene needed to be expressed at 30 arbitrary expression units in the trichome sample in order be called a trichome gene by Jakoby et al. (2008)). As the data includes expression in *gl3* and *try* mutants, I looked for genes that responded to these mutations in similar ways as GL2. GL2 expression is reduced in *gl3* mutant and increased in *try* mutant compared to wild type with the difference between these two ratios for GL2 being 0.73. Therefore the list of genes with high trichome and low pavement expression was sorted for genes that had a difference between the *gl3*/wild type ratio and the *try*/wild type ration of at least 0.5. Meaning that a gene passing this test would have twice the relative expression in *try* mutants as it has in *gl3* mutants. This analysis left 38 genes out of 3,231 in the data set, which contains only the genes that were found to have increased trichome expression compared to the leaf tissue with its trichomes removed (Jakoby et al. 2008). . Interestingly GL2 is included in these 38 genes. Then to further examine the probability that these genes were GL2 targets, the list was sent through Promomer to check for the presence of an L1-box. 31 of the 38 genes had at least the minimum required L1-box. The last step of the analysis of these genes was to determine if the gene is expressed in qualitatively suggestive pattern using AtGenExpress. Through all this analysis 14 genes were selected for further analysis (Table 4.2).

Candidate expression in *gl2* mutant

The first experimental test of candidate genes identified through correlation and microarray analysis, was to determine if their expression changes in the absence of GL2. This was accomplished with real time PCR of wild type (L*er*) and *gl2* mutant (*gl2-1*)

seedlings with fully developed first true leaves and expanding second leaves. In addition to 22 genes from the screens, a set of known trichome genes with L1-boxes were also included in the experiment. 8 of the 27 tested genes have virtually no expression change in *gl2* mutant compared to wild type (Fig 4.1). 11 were downregulated in *gl2* mutant approximately 2 to 4 fold (Fig 4.1). The remaining 8 genes were downregulated greater than 4 fold in *gl2* mutant (Fig 4.1).

Promoter GUS analysis of candidate genes

The eight most downregulated candidate genes in *gl2* mutant received the next layer of experimental analysis which was examining their developmental expression pattern to verify expression in trichomes, or at least in the same tissues as GL2. Promoter:GUS constructs were created and transformed into Columbia (Col) wild type and *gl2* mutant lines. Each promoter consisted of at least 2 kb of upstream sequence from the start codon.

At5g15160 is a bHLH protein. Based on its GUS expression pattern it is not a trichome gene and does not look like a target of GL2. It is expressed at the tips of the serrated teeth of the leaves including the tip of the leaf (Fig 4.3A and C). In Arabidopsis the leaf has at least 1 set of serrated points on the first true leaves with more on older leaves. At5g15160 is not expressed in seeds but has expression in the root tip at what appears to be the quiescent center (Fig 4.3B).

At1g06100, a fatty acid desaturase, has a trichome specific expression pattern with levels decreasing as the trichome matures, though expression never completely ceases (Fig 4.2A). It also has expression in the chalazal end of the seed coat (Fig 4.2C).

In the root there is expression at the branch points of lateral roots that persists well after the root elongates (Fig 4.2B). In *gl2-1*, At1g06100 still is expressed in the trichomes that develop, however expression in the root branch points disappears.

At4g17860, an unknown protein, has expression in trichomes specifically that reduces as the leaf matures (Fig 4.2D). It also is expressed in the radicle of the developing embryo (Fig 4.2F) and in the root tip (Fig 4.2E). In *gl2-1*, at4g17860 is still expressed in trichomes and root tips.

At5g33370, a GDSL-motif lipase/hydrolase, is expressed specifically in trichomes (Fig 4.3D). It is also expressed in root tips (Fig 4.3E). In *gl2-1*, At5g33370 is still expressed in trichomes however in is no longer expressed in roots.

At1g01600 or CYP86A4, a predicted fatty acid hydroxylase, is expressed in trichomes and in a diffuse pattern at the base of the developing leaf (Fig 4.2G). It is also expressed in the root including root hairs (Fig 4.2H). Inside developing siliques there is expression in the funiculus leading to the ovules (Fig 4.2I).

At3g61840, an unknown protein, is expressed in developing trichomes, however expression ceases as the leaf matures (Fig 4.3G). There is also expression in root tips (Fig 4.3H). In *gl2-1*, At3g61840 is no longer expressed in roots, but expression remains in trichomes though expression does not decrease as trichomes mature at the same rate as in wild type.

At3g18180, a glycosyl transferase, has no expression in trichomes, leaves, roots or seeds in wild type. However in *gl2-5* seeds, there is seed coat expression (Fig 4.3F).

At5g65300, an unknown protein, has trichome specific expression (Fig 4.4A and B). It is also expressed in root tips in what appears to be the non-hair files (Fig 4.4C). In *gl2-1* there is no longer any expression in trichomes or roots.

ChIP analysis of targets

ChIP assays were used to verify if the possible targets of GL2 are direct targets. Since GL2 overexpression caused dominant negative phenotypes I was forced to use an HA-tagged GL2 construct under control of its own promoter. This resulted in a greatly reduced amount of tagged protein in the experiment. At5g65300 and EGL3 showed strong positive results in a set of experiments (Fig 4.5). While TRY, At1g01600 and At1g06100 revealed weaker though likely positive results (Fig 4.6). At5g33370, At4g17680, At3g61840, and PLDZ, the reported root target, had inconclusive results.

Differential sequence analysis of GL2 and TTG2

SOLiD next generation sequence technology allowed large scale differential sequencing analysis for genes whose expression changes in *gl2* and *ttg2* mutants compared to wild type. SOLiD technology can sequence hundreds of millions of reads of 35 or 50 bp at a time. Using this technology, I was able to produce approximately 20 million sequence reads for each of wild type, *gl2* mutant, and *ttg2* mutant cDNA. Of these 20 million reads, 40-70% could be mapped to the *Arabidopsis* genome with no more than 3 mismatches. To be able to compare the relative expression levels of each gene across the data sets the number of reads per gene was normalized to the total number of mapped reads per data set. The normalized gene counts allowed me to determine the fold change for each gene between wild type and mutant.

Of the 22,057 genes sequenced for the *gl2-5*/Col experiments, which had at least some expression in both mutant and wild type, 749 genes were down-regulated in *gl2-5* mutant at least 2 fold. An additional 1,219 genes were expressed in wild type but not in gl2 mutant though of these only 159 have enough reads to suggest at least 2 fold change in expression. Of the 749 genes down-regulated in *gl2-5,* 683 have the minimum sequence required for an L1-box while 214 have a complete L1-box sequence. 961 genes with expression in both samples are up-regulated in gl2 mutant at least 2 fold. Of these 847 have a minimum L1-box and 251 have a complete L1-box. There are an additional 598 genes that had expression in *gl2-5* but not in wild type, however only 154 have expression suggesting at least a 2 fold change.

Over representation analysis (ORA) of the up and downregulated gene sets using Genetrail yielded some insight into what is GL2's function. Both data sets had genes involved with the endomembrane system and transcription over represented compared to the total set of genes sequenced. The downregulated gene set (Table 4.3) had many more types of GO categories related to transcriptional regulation, fatty acid hydroxylation, inositol oxygenase, and palmitoyl-(protein) hydrolase activity, while the upregulated data set (Table 4.4) had genes involved in cell recognition, transporters and lipid transport over represented.

In the *ttg2-1*/L*er* data set 22,294 genes were sequenced. Of these 980 genes were downregulated at least 2 fold in *ttg2* mutant compared to wild type. 705 of those had the minimum sequence necessary for a W-box which is the predicted DNA binding element for TTG2. There were also 438 genes with no expression in *ttg2* but expression in wild type suggesting at least 2 fold change. The genes upregulated in *ttg2* mutant consisted of

3,012 genes with at least 2 fold change, of which 2,211 had a minimum W-box. An additional 675 genes with expression in *ttg2* and no expression in wild type suggesting at least 2 fold change. ORA analysis of the downregulated data set (Table 4.5) revealed an overrepresentation of genes involved in cell wall modification/organization and genes associated with the cell membrane. Along those same lines there was an enrichment of a GO category associated with exocytic vesicles. Another interesting enriched GO category is that of release from seed dormancy. ORA analysis of the TTG2 upregulated data set (Table 4.6) using Genetrail had an over representation of processes involved in cell signaling.

Analysis of the expression of known trichome genes in the GL2 and TTG2 SOLiD data revealed that most were unchanged in the mutant compared to wild type. *MYB23, MYB5, CPC, NOK,* and *TCL1* were downregulated at least two fold in *gl2* mutant while *CPC, MYB23,* and *TCL1* were downregulated at least two fold in *ttg2-1* (Table 4.7). *GL3, ETC2*, and *ETC3* were upregulated at least two fold in *gl2* mutant while *TRY, ETC1, CPR5, WRM, DIS1, BLT, KIC,* and *MYB5* were upregulated at least two fold in *ttg2-1* (Table 4.7).

Comparison of the genes with substantial change in the data sets allows us to examine the possible overlapping function of these two transcription factors (Fig 4.7). 136 genes were downregulated and 432 genes upregulated at least 2 fold in both *gl2-5* and *ttg2-1.* It is possible that these genes are trichome genes as they require both *gl2* and *ttg2* for proper expression. There are also a number of genes that are down in *ttg2-1* while up in *gl2-5* and vice versa. These genes could be trichome genes under regulation of GL2 and/or TTG2 but in different ways.

Mutant Analysis

Genes identified as possible GL2 targets from the various screens discussed were analyzed for mutant phenotypes using T-DNA insertion lines available from ABRC. None of the potential target insertion lines from the correlation studies or the trichome microarray had any visible trichome or seed phenotype.

100 possible targets of GL2 and 127 possible targets of TTG2 from the SOLiD analysis were screened for trichome mutations however no visible phenotypes have been identified.

Overexpression Analysis of Possible Targets

Each of the 8 possible targets identified through the correlation and microarray analysis was overexpressed in both wild type and *gl2* mutant lines. However other than At5g65300 which will be discussed in the next chapter no visible trichome phenotype was identified.

Overexpression of At5g15160 does have a phenotype in wild type and *gl2-1* mutants. Transgenic plants display greatly elongated hypocotyls (Fig 4.8). Otherwise the plants appear normal.

MATERIALS AND METHODS

Bioinformatic Analysis

Correlation analysis was performed using on line Expression Angler software (http://bar.utoronto.ca/ntools/cgi-bin/ntools_expression_angler.cgi) from BAR, with the setting to retrieve the top 100 most correlated genes in each of: AtGenExpress Consortium's Tissue series; Botany Array Resource; and NASCArrays data sets.

Promoter analysis was performed using Promomer (http://bar.utoronto.ca/ntools/cgi-bin/BAR_Promomer.cgi) from BAR using the fourth search option to find binding elements in promoters of a list of genes. Promoter regions from each of the potential targets were searched for L1-boxes TAAATG(T/C)A, W-Boxes (T)(T)TGAC(T/C) and the mininium required sequences for each.

Qualitative analysis of expression of possible targets was checked with the AtGenExpress Visualization tool (http://jsp.weigelworld.org/expviz/expviz.jsp).

Over-representation analysis was performed using Genetrail (http://genetrail.bioinf.uni-sb.de/index.php).

SOLiD Analysis

Two sets of 0.25g each of 10 day old *Arabidopsis* plants were harvested with cotyledons, hypocotyls, and roots removed leaving the first set of true leaves, the just expanding second pair of true leaves, and the meristem from each of Col, L*er*, *ttg2-1*, and *gl2-1*. Total RNA was extracted from each sample using Qiagen RNeasy plant Mini Kit. Then mRNA was isolated from each sample using Qiagen Oligotex mRNA Kit. The mRNA was used as template for Invitrogen's SuperScript Double-Stranded cDNA Synthesis Kit using oligo dT. Double stranded cDNA was then sent to the University of Texas at Austin's Genome Sequencing and Analysis Facility (GSAF) for further processing, library creation, and sequencing. Each sample was sequencing using 1/32nd of SOLiD run which should correspond to at least 10 million reads per sample.

The sequence reads were mapped to the TAIR9 release reference genome using Mapreads from ABI allowing no more than 3 mismatches. Reads that mapped to

multiple positions were removed from the data set. Then the mapped positions of the reads were translated into their corresponding genes and the number of reads per gene counted. The read counts per gene were normalized to the total number of mapped reads for that sample.

Gene expression analyses

Total RNA was prepared from L*er* wild type and *gl2-1* 10 day old plants using Qiagen RNeasy plant Mini Kit. 1 ug of RNA was used in 20 ul reverse transcription reactions containing 250 nM Actin and oligo dT primers. Parallel 25 ul PCR reactions were prepared using cDNA reactions as templates with half volume of 2X SuperPower Syber mixture (ABI) and run on a spectrofluorometric thermal cycler (ABI 7900HT). For each target, five PCR reactions containing 400 nM primers and 1 ul first strand cDNA as template were performed alongside four actin control PCR reactions containing 200 nM Actin primers and 1 ul first strand cDNA. The comparative cycle threshold method was used to analyze the results (User Bulletin 2, ABI PRISM Sequence Detection System). Each experiment was performed twice for each target with consistent results. Results of representative experiments are presented. Primers used for each target included in Table 4.10.

Promoter GUS Analysis

At least 2kb of upstream promoter from At5g65300, At1g01600, At1g06100, At3g61840, At5g33370, At4g17860, At3g18180, and At5g15160 was cloned into pBGWFS7 (Karimi et al 2002) using Invitrogen's Gateway system. Primers for each

listed in Table 4.12. Each was transformed into at least wild type Col and *gl2-1* using Agrobacteria floral dip transformations.

Analysis of promoter GUS transgenic lines was performed with various tissues stained with X-Gluc as previously described (Masucci et al., 1996). Staining was visualized using Nikon dissecting scope.

Chromatin Immunoprecipitation (ChIP) Experiments

ChIP experiments were performed as described previously (Morohasi et al., 2007). The tissue used was from *gl2-1* plants expressing GL2-HA under control of its own promoter. This construct was created using Gateway. 2kb of upstream promoter with the complete GL2 gene with no stop codon from Col genomic was cloned into pEarleyGate301. PCR was performed with the primers in Table 4.11.

Mutant Lines

T-DNA lines for each possible target listed in tables 4.1, 4.2, 4.8, and 4.9 were ordered from ABRC.

Overexpression Analysis

Coding region from start to stop for each of At5g65300, At1g01600, At1g06100, At3g61840, At5g33370, At4g17860, At3g18180, and At5g15160 was cloned into pB7WG2 using Gateway. Primers sequences can be found in table 4.13. Each was transformed into at least wild type Col and *gl2-5* using Agrobacteria floral dip transformations.

DISCUSSION

Identifying targets of GL2 and TTG2 is vital for the understanding of what is required for an epidermal cell to become a properly formed trichome cell. Currently there are no known targets for either gene in the trichome development pathway, although GL2 has 3 published direct targets in root hair development. Using bioinformatic analysis I was able to identify several new targets of GL2. In addition, the expression data obtained through SOLiD sequencing resulted in a new database that will be a great source of additional targets. While for the most part, no phenotypes were identified for any of the new targets, they are trichome genes downstream of GL2. Several seem to have possible homologs that could require multiple mutants for phenotypes to appear.

An interesting set of GL2 targets is the TTG1 dependent transcription factor complex members, EGL3, MYB23, and MYB5. These complex members have been shown to be specialized for the maturation of the trichome, especially branching. This was observed in their mutants, which have slight defects in branching. Also, overexpression experiments with GL3 and EGL3 in *ttg1* mutant shows that EGL3 produces a moderate level of normal looking trichomes while GL3 produces many more trichomes of odd shapes and branch patterns (Zhang et al. 2003). Suggesting EGL3 is much better at the maturation of trichomes than its homolog GL3, while GL3 is better at initiation. MYB23 also has other experimental evidence showing its specialization in the maturation stage of trichome development. In complementation experiments, GL1, the major trichome initiation myb, cannot totally replace MYB23 function (Kirik et al. 2005). There is further evidence that MYB23, MYB5, and EGL3 are later acting in the

comparison of the expression patterns of these genes and their homologs, GL1 and GL3. All three are expressed in trichome cells indefinitely while GL1 and GL3 both stop being expressed as the trichome develops (Zhao et al. 2008; Kirik et al. 2005; Gonzales et al. 2009). This specialization of complex members allows the complex to be able to regulate development across the entire life cycle of the trichome with different sets of targets for different processes. Initiation, branching, wall thickening, and papillae formation all could be under the regulation of the complex yet there are likely different sets of genes required for each process. With all 3 of the late acting complex members being transcriptional targets of GL2, while at the same time GL2 is a target of the complex, results in an interesting feedback loop or switch. I suggest that GL1/GL3/TTG1 controls initiation and starts activation of GL2. GL2 then goes on to control elongation of the trichome and activates EGL3, MYB23, and MYB5, which together with TTG1 facilitate proper branching and maturation of the trichome. GL2 as a bridge between trichome intiation and maturation may explain why overexpression of GL2 seems to upset the balance resulting in a dominant negative phenotype where trichomes are prevented from maturing.

GL2 seems to have a role with lipids. One of its known targets is a phospholipase D (Ohashi et al. 2003) and 3 of the novel targets identified here are lipid modifying enzymes. In addition, the SOLiD sequencing showed enrichment of fatty acid hydroxylase, inositol oxygenase, and palmitoyl-(protein) hydrolase activity. These results, along with gl2's increased seed oil phenotype, and the sterol binding START domain, are strong evidence of GL2 acting through lipid manipulations somehow. The predicted role of phospholipase D is to produce signals for vesicle targeting within the

cell. This could be important in moving cell components to the proper position in the elongating trichome. GL2's role as the major regulator of elongation of the trichome could require lipid modification for the growing membrane. There is also the possibility of GL2 controlled lipids being required as signal molecules for transcription factors, such as GL2 itself. GL2, with its START domain that likely regulates its own function in some way, could be caught in a feedback loop with its targets. This may be to maintain uniform elongation throughout development so that each trichome elongates relatively the same amount.

All of these conclusions can be combined to give a new picture of how GL2 functions. Through its targets, GL2 acts as a switch between the initiation complex and the maturation complex. GL2 has many roles with lipids that are likely involved in the elongation process within the trichome, in addition to creating a likely feedback loop with GL2. There is one more role that will be described in further detail in the following chapter that places GL2 in control of accessory cell fate though its target At5g65300 (Fig 4.9).

TTG2 mutants seem to have major defects in trichome branching, maturation and morphology. SOLiD differential sequence analysis found several known cytoskeleton genes involved in trichome branching with substantial expression changes. As *ttg2* mutants have phenotypes similar to many of the cytoskeleton mutants these could be likely targets of TTG2. TTG2 mutant trichomes seem to have reduced cell wall thickening which was also born out in the SOLiD data. Genes involved in wall structure were enriched in the downregulated gene set verifying TTG2's possible role with the cell wall. These genes are also very likely targets of TTG2.

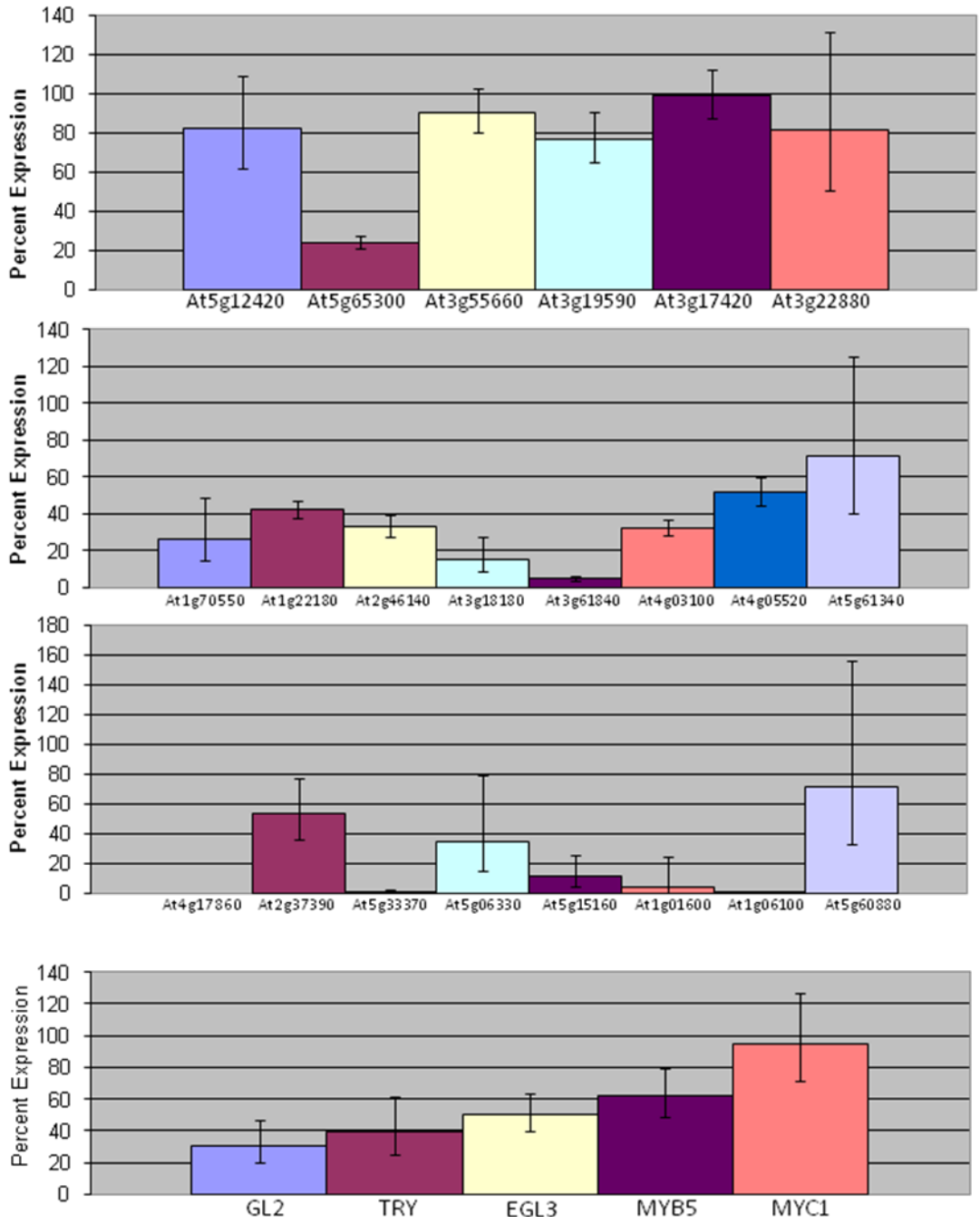

Fig 4.1. RT-PCR Expression Analysis of Possible GL2 Targets.

Results displays expression in *gl2-1* mutants as a percent of expression in L*er* wild type.

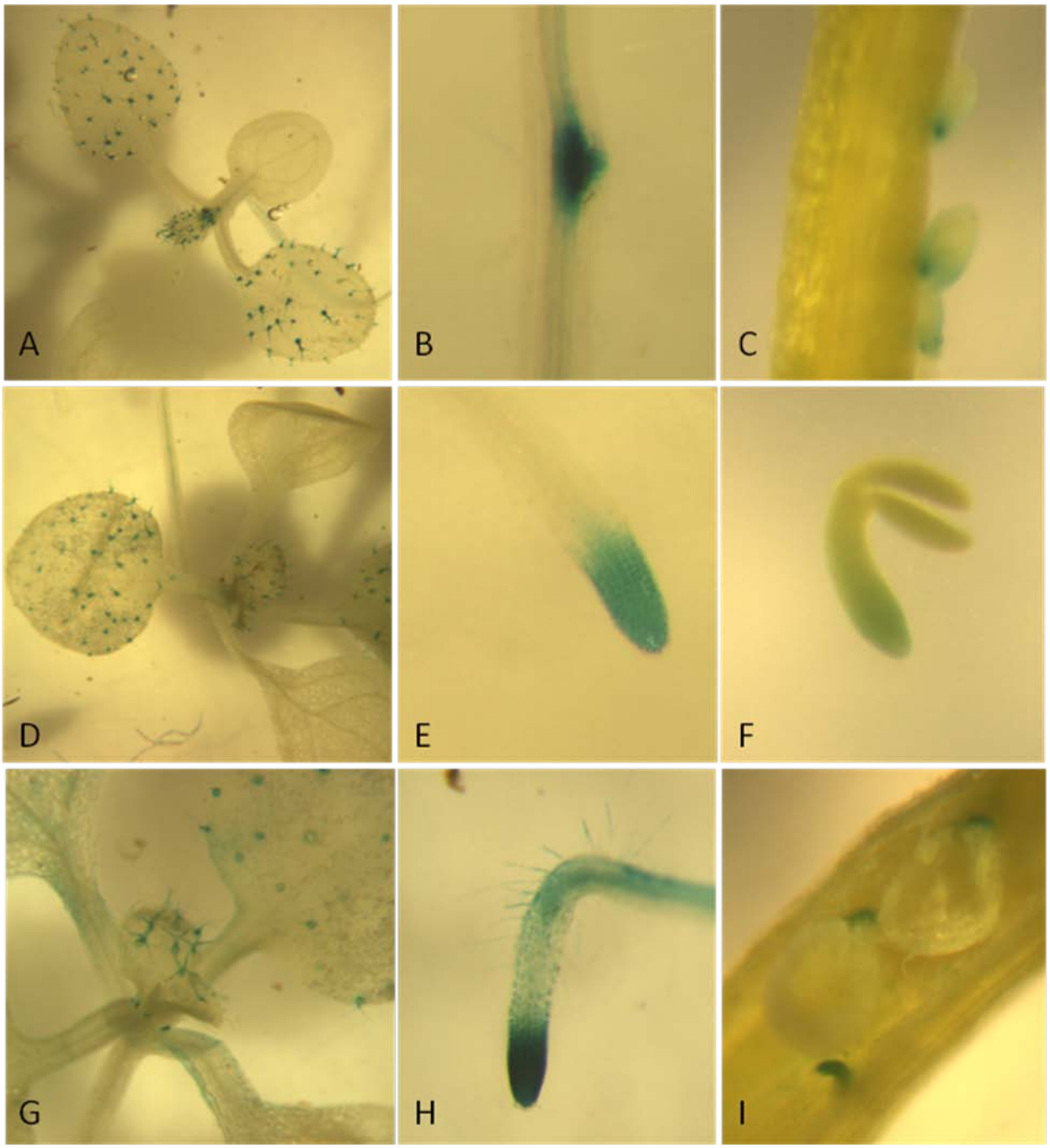

Fig 4.2. Promoter GUS Analysis of Possible GL2 Targets.

Promoter GUS expression of possible targets in 12 days old shoots, roots and seeds. (A-C.) At1g06100, (D-F.) At4g17860, (G-I.) At1g01600 in Col ecotype.

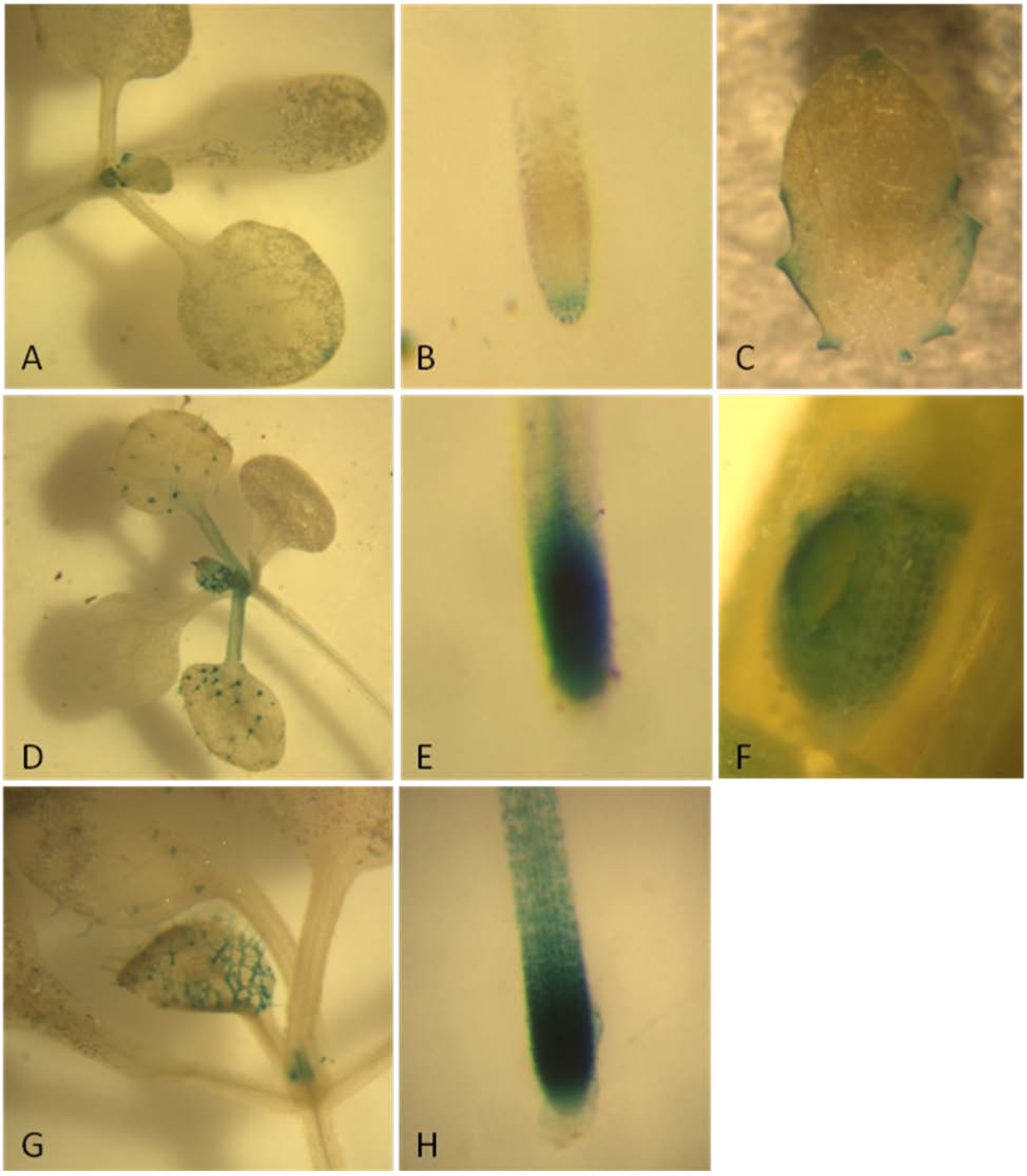

Fig 4.3. Promoter GUS Analysis of Possible GL2 Targets.

Promoter GUS expression of possible targets in 12 days old shoots, roots and seeds. (A-C.) At1g15160, (D-E.) At5g33370, (G-H.) At3g61840 in Col ecotype . Expression of (F.) At3g18180 in *gl2-5*.

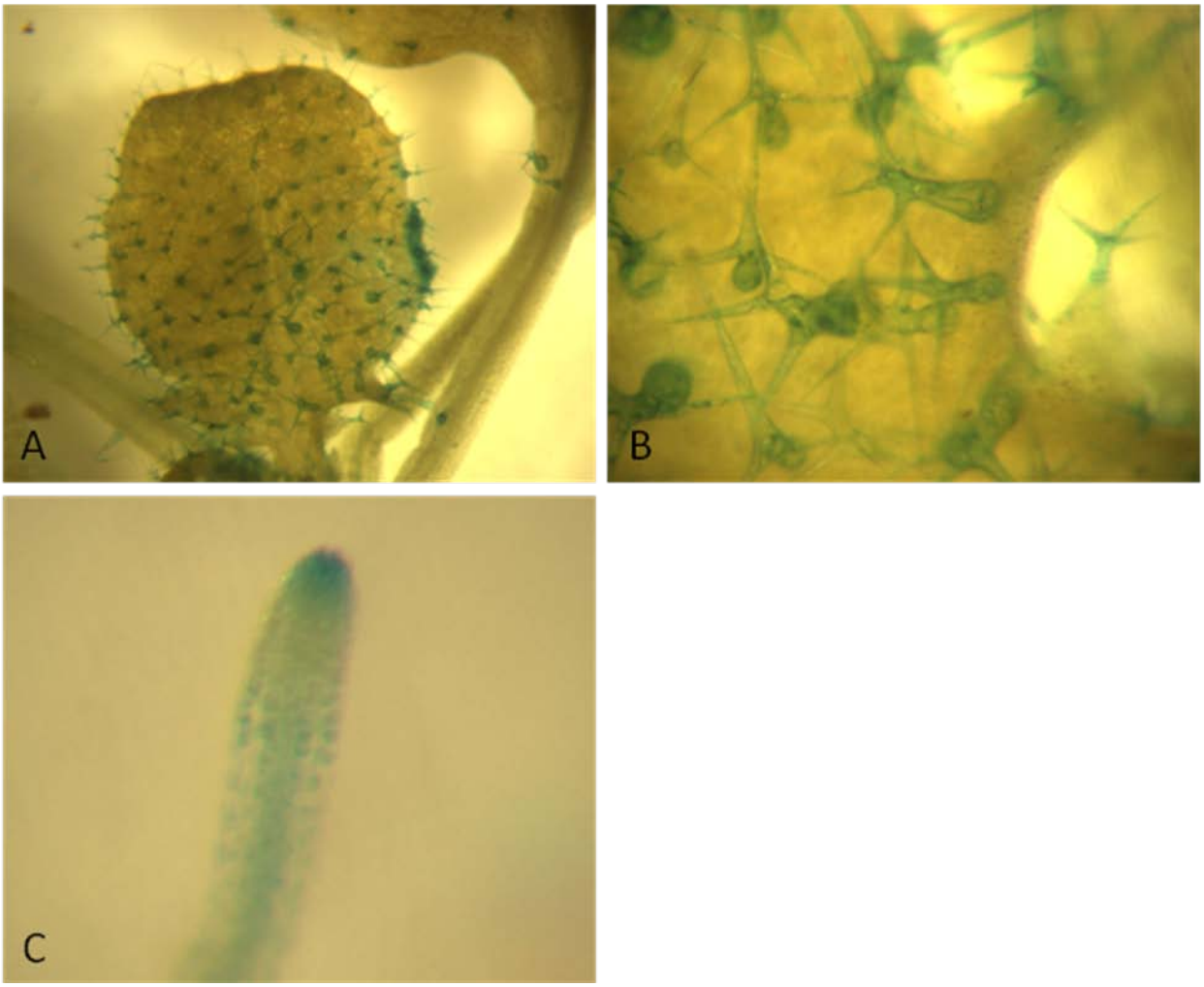

Fig 4.4. Promoter GUS Analysis of Possible GL2 Targets.

Promoter GUS expression of At5g65300 in 12 days old shoots and roots

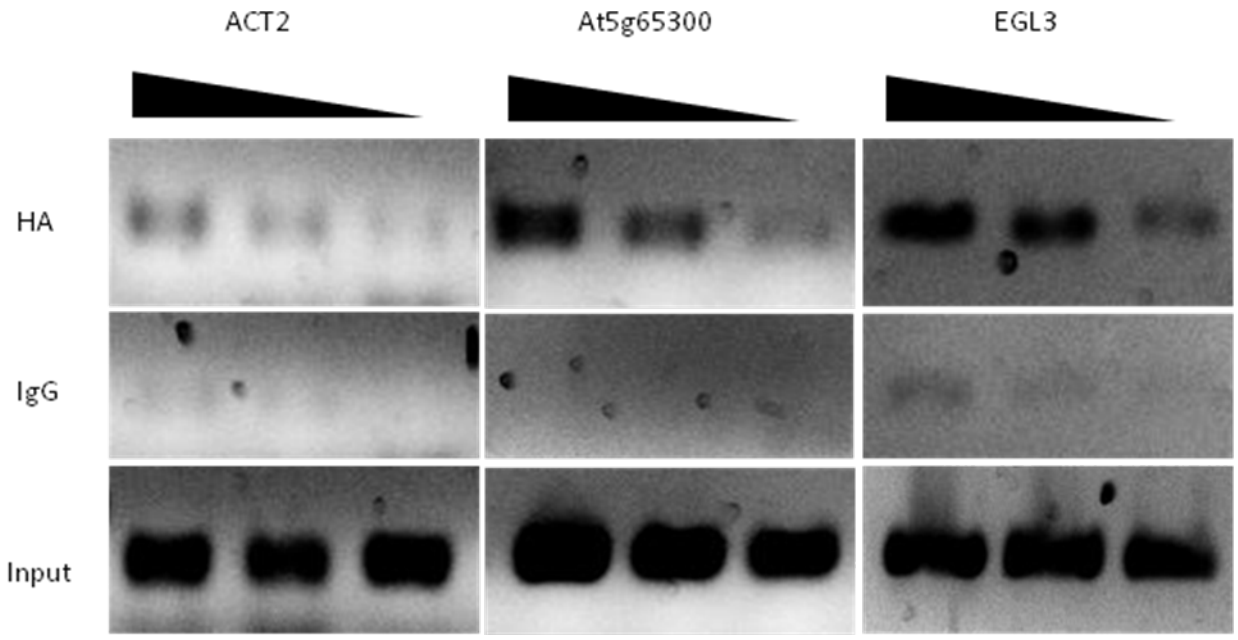

Fig 4.5. GL2 ChIP-PCR of Possible Targets.

ChIP using GL2::GL2-HA transgenic plants with PCR of ACT2, At5g65300, and EGL3. Triangles represent 2 fold serial dilution of elution from HA antibody on affinity beads.

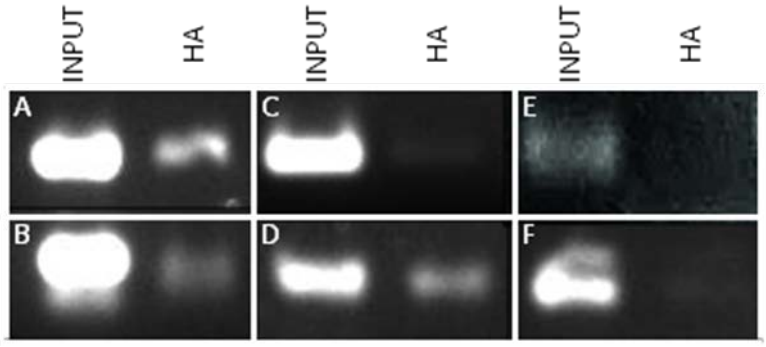

Fig 4.6. GL2 ChIP-PCR of Possible Targets.

ChIP using GL2::GL2-HA transgenic plants with PCR of (A.) EGL3, (B.) At5g65300, (C.) At1g01600, (D.) At1g06100, (E.) ACT2, and (F.) TRY.

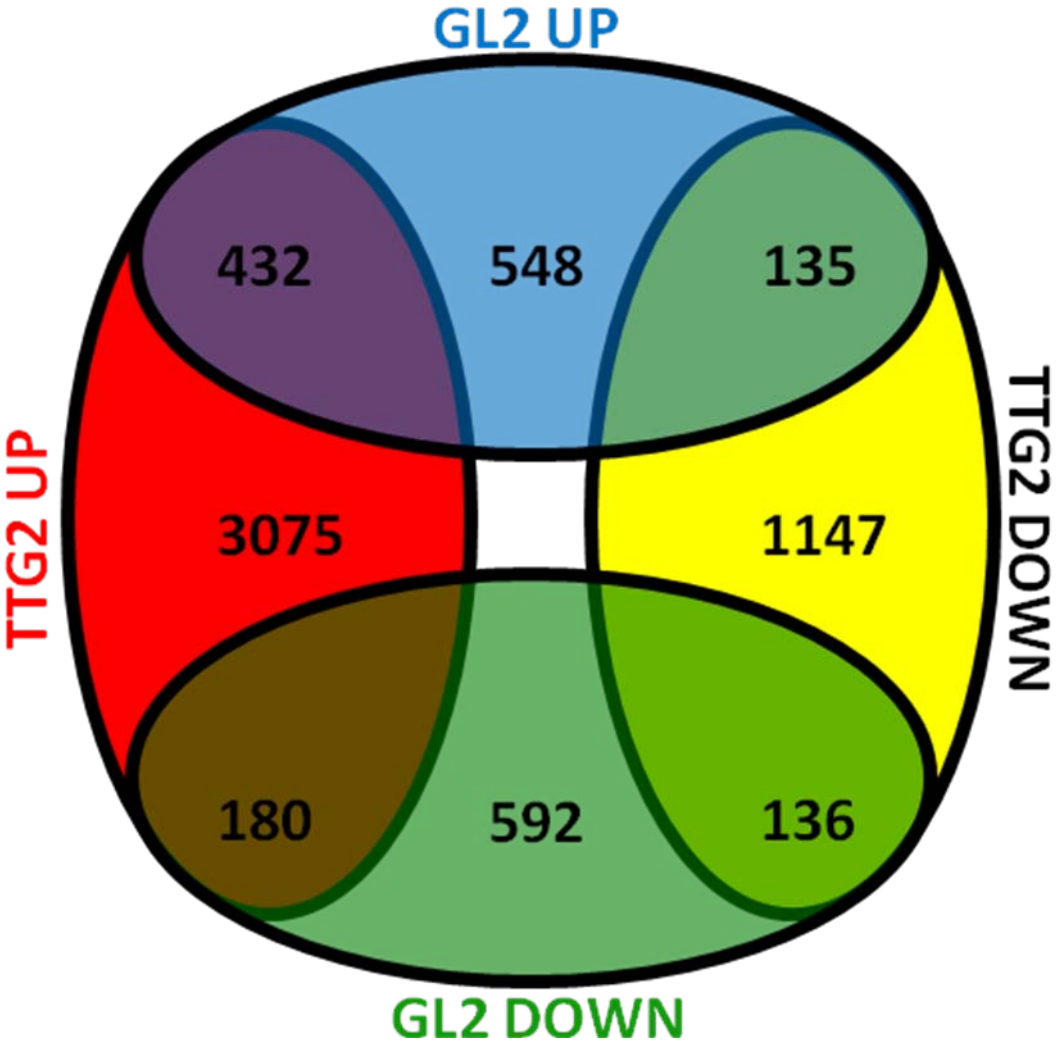

Fig 4.7. Genes in Common Between TTG2 and GL2 Solid Data Sets.

Venn diagram of genes with at least 2 fold change compared to wild type in *gl2* and *ttg2* data sets. UP refers to genes with at least 2 fold increase in expression in mutant sample compared to wild type and DOWN refers to genes with at least 2 fold decrease in expression in mutant sample compared to wild type.

Fig 4.8. Overexpression of At1g15160 in Columbia Wild Type.

Primary transgenic plants overexpressing At1g15160 in Col grown in a field of untransformed wild type.

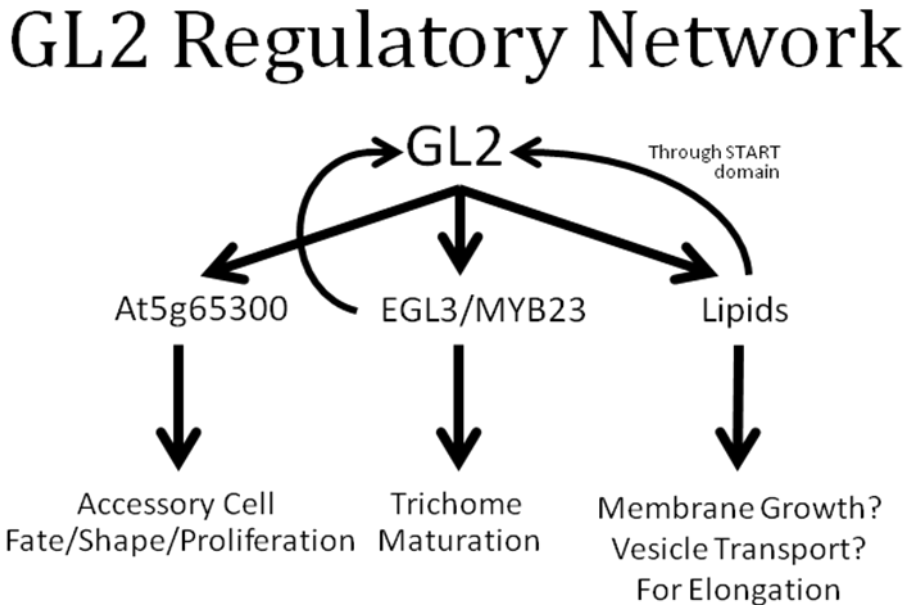

Fig 4.9. Model of GL2 Regulatory Network.

Table 4.1 Possible GL2 Targets Identified Through Correlation Studies

	Pearson Coefficient	L1-box	T-DNA		
AT2G37390	0.768	Y	SALK_086647		
AT3G19590	0.719	Y	SALK_028505C	SALK_148060C	
AT3G22880	0.728	Y	SALK_056177		
At4g05520	0.766	Y	SALK_128099C		
At5g60880	0.684	Y	SALK_086936C		
At3g20150	0.616	Y			
At4g03100	0.668	Y			
At3g55660	0.698	Y			
AT3G18180	0.893	Y	SALK_047132	SALK_047126	SALK_057802
At4g09820	0.941	Y	TT8		
At3g13540	0.883	Y	MYB5		
At5g40330	0.644	Y	MYB23		

Table 4.2 Possible GL2 Targets Identified through Trichome Database

Gene ID	T-DNA
AT5G06330	SALK_135329C
AT1G06100	SALK_070278C, SALK_009538C, SALK_048849
AT1G22180	SALK_058361C
AT2G46140	SALK_146406, SALK_106690C
AT3G17420	SALK_047485C
AT3G61840	SALK_024950C, SALK_043757, SALK_043871, SALK_044095
AT4G17860	SALK_100404, SALK_019862
AT5G12420	SALK_028121, SALK_062207
AT5G15160	CS122628, SALK_135685
AT5G33370	SALK_032531C, SALK_070013, SALK_005902, SALK_037114
AT5G65300	SALK_069313C, SALK_133788C
AT1G01600	SALK_077857C, SALK_013462, SALK_073079, SALK_085358, SALK_013462
AT5G61340	SALK_062281
AT1G70550	SALK_060907C

Table 4.3 Enrichment Analysis of Genes Downregulated in *gl2-1* Compared to Col

Category	Subcategory	Subcategory alternative name	expected	observed	p-value
GO	transcription factor activity	GO:0003700	53.8515	93	8.64E-06
GO	transcription regulator activity	GO:0030528	61.4294	97	0.000173
GO	regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	GO:0019219	51.7152	84	0.000256
GO	regulation of transcription	GO:0045449	50.4253	82	0.000305

GO	regulation of nitrogen compound metabolic process	GO:0051171	52.4004	84	0.000374
GO	regulation of macromolecule biosynthetic process	GO:0010556	51.957	83	0.000433
GO	regulation of biosynthetic process	GO:0009889	53.2469	84	0.000552
GO	regulation of cellular biosynthetic process	GO:0031326	53.2469	84	0.000552
GO	regulation of primary metabolic process	GO:0080090	54.7786	86	0.000552
GO	palmitoyl-(protein) hydrolase activity	GO:0008474	0.241848	4	0.00091
GO	regulation of cellular metabolic process	GO:0031323	56.9955	87	0.001191
GO	transcription	GO:0006350	53.8112	83	0.001219
GO	rRNA modification	GO:0000154	2.29756	10	0.00163
GO	regulation of macromolecule metabolic process	GO:0060255	57.7614	87	0.00163
GO	regulation of transcription, DNA-dependent	GO:0006355	26.3211	47	0.001816
GO	oxygen binding	GO:0019825	6.57021	18	0.00195
GO	regulation of gene expression	GO:0010468	55.9072	84	0.002122
GO	regulation of RNA metabolic process	GO:0051252	26.6033	47	0.002142
GO	regulation of metabolic process	GO:0019222	61.8728	90	0.003477
GO	iron ion binding	GO:0005506	9.75454	22	0.00504
GO	transcription, DNA-dependent	GO:0006351	27.9335	47	0.005297
GO	RNA biosynthetic process	GO:0032774	28.0141	47	0.005461

GO	DNA binding	GO:0003677	76.3031	105	0.006311
GO	endomembrane system	GO:0012505	96.5377	127	0.009219
GO	fatty acid (omega-1)-hydroxylase activity	GO:0008393	0.241848	3	0.014915
GO	RNA modification	GO:0009451	4.31296	12	0.01554
GO	small nucleolar ribonucleoprotein complex	GO:0005732	3.22464	10	0.016518
GO	heme binding	GO:0020037	7.5376	17	0.018156
GO	rRNA processing	GO:0006364	3.42618	10	0.024843
GO	rRNA metabolic process	GO:0016072	3.42618	10	0.024843
GO	tetrapyrrole binding	GO:0046906	8.58561	18	0.028268
GO	response to chitin	GO:0010200	4.91758	12	0.040651
GO	zinc ion transmembrane transporter activity	GO:0005385	0.362772	3	0.04536
GO	zinc ion transport	GO:0006829	0.362772	3	0.04536
GO	inositol oxygenase activity	GO:0050113	0.120924	2	0.04652

Table 4.4 Enrichment Analysis of Genes Upregulated in *gl2-1* Compared to Col.

Category	Subcategory	Subcategory alternative name	expected	observed	p-value
KEGG	Naphthalene and anthracene degradation	626	1.25187	9	3.17E-05
KEGG	Stilbenoid, diarylheptanoid and gingerol biosynthesis	945	1.33178	9	3.17E-05
KEGG	gamma-Hexachlorocyclohexane degradation	361	1.30514	9	3.17E-05

KEGG	Limonene and pinene degradation	903	1.41168	9	3.95E-05
KEGG	Diterpenoid biosynthesis	904	0.319626	3	0.016622
GO	endomembrane system	GO:0012505	115.208	211	3.21E-17
GO	cellular calcium ion homeostasis	GO:0006874	1.10638	7	0.001489
GO	calcium ion homeostasis	GO:0055074	1.10638	7	0.001489
GO	defense response	GO:0006952	30.7381	52	0.002934
GO	ligand-gated ion channel activity	GO:0015276	1.25069	7	0.00307
GO	ligand-gated channel activity	GO:0022834	1.25069	7	0.00307
GO	intracellular ligand-gated ion channel activity	GO:0005217	0.962069	6	0.004195
GO	di-, tri-valent inorganic cation homeostasis	GO:0055066	2.45328	9	0.009318
GO	cellular di-, tri-valent inorganic cation homeostasis	GO:0030005	2.06845	8	0.012218
GO	transcription factor activity	GO:0003700	64.2662	89	0.014137
GO	regulation of transcription	GO:0045449	60.1774	84	0.015231
GO	transcription regulator activity	GO:0030528	73.3096	99	0.015989
GO	cellular metal ion homeostasis	GO:0006875	1.87603	7	0.022749
GO	oxygen binding	GO:0019825	7.84086	17	0.022749
GO	metal ion homeostasis	GO:0055065	1.87603	7	0.022749
GO	cation homeostasis	GO:0055080	3.41534	10	0.022749
GO	regulation of nitrogen compound metabolic process	GO:0051171	62.5345	85	0.023027

GO	regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	GO:0019219	61.7167	84	0.023211
GO	regulation of macromolecule biosynthetic process	GO:0010556	62.0053	84	0.024971
GO	gated channel activity	GO:0022836	2.50138	8	0.026591
GO	iron ion binding	GO:0005506	11.641	22	0.026795
GO	regulation of primary metabolic process	GO:0080090	65.3726	87	0.02969
GO	regulation of macromolecule metabolic process	GO:0060255	68.9322	91	0.029876
GO	heme binding	GO:0020037	8.99534	18	0.030763
GO	regulation of gene expression	GO:0010468	66.7195	88	0.033631
GO	toxin metabolic process	GO:0009404	2.11655	7	0.033836
GO	toxin catabolic process	GO:0009407	2.11655	7	0.033836
GO	transcription	GO:0006350	64.2181	85	0.034549
GO	cell recognition	GO:0008037	1.15448	5	0.034549
GO	recognition of pollen	GO:0048544	1.15448	5	0.034549
GO	ion homeostasis	GO:0050801	3.84827	10	0.034549
GO	regulation of biosynthetic process	GO:0009889	63.5446	84	0.035398
GO	regulation of cellular biosynthetic process	GO:0031326	63.5446	84	0.035398
GO	substrate-specific channel activity	GO:0022838	5.14707	12	0.035442
GO	channel activity	GO:0015267	5.19517	12	0.036851
GO	passive transmembrane transporter activity	GO:0022803	5.19517	12	0.036851
GO	lipid localization	GO:00108	4.5698	11	0.03

		76	3		7477
GO	glutathione transferase activity	GO:0004364	2.21276	7	0.038266
GO	quercetin 3'-O-glucosyltransferase activity	GO:0080045	0.14431	2	0.040083
GO	systemic acquired resistance	GO:0009627	1.77983	6	0.045929
GO	cellular cation homeostasis	GO:0030003	2.93431	8	0.04927
GO	lipid transport	GO:0006869	4.185	10	0.049346
GO	innate immune response	GO:0045087	12.6993	22	0.049346

Table 4.5 Enrichment Analysis of Genes Down regulated in *ttg2-1* compared to L*er*.

Category	Subcategory	Subcategory alternative name	expected	observed	p-value
KEGG	Photosynthesis - antenna proteins	196	0.725259	8	6.17E-06
KEGG	Naphthalene and anthracene degradation	626	1.90858	9	0.001115
KEGG	Stilbenoid, diarylheptanoid and gingerol biosynthesis	945	2.02309	9	0.001115
KEGG	gamma-Hexachlorocyclohexane degradation	361	1.98492	9	0.001115
KEGG	Limonene and pinene degradation	903	2.13761	9	0.001389
KEGG	Carotenoid biosynthesis	906	0.725259	4	0.029583
GO	endomembrane system	GO:0012505	142.833	207	1.14E-06
GO	oxygen binding	GO:0019825	9.89811	30	3.21E-06
GO	anchored to membrane	GO:0031225	11.5778	33	3.21E-06
GO	rRNA modification	GO:0000154	3.59931	16	1.31E-05
GO	tetrapyrrole binding	GO:0046906	12.4176	32	2.54E-05
GO	monooxygenase activity	GO:0004497	12.7176	32	3.88E-05

GO	small nucleolar ribonucleoprotein complex	GO:0005732	4.97905	17	0.000169
GO	electron carrier activity	GO:0009055	20.3961	42	0.000169
GO	chlorophyll binding	GO:0016168	1.49971	9	0.00019
GO	(+)-abscisic acid 8'-hydroxylase activity	GO:0010295	0.239954	4	0.000283
GO	rRNA processing	GO:0006364	5.21901	17	0.000284
GO	rRNA metabolic process	GO:0016072	5.21901	17	0.000284
GO	regulation of transcription	GO:0045449	74.1459	108	0.000869
GO	transcription factor activity	GO:0003700	78.885	113	0.0011
GO	carboxylesterase activity	GO:0004091	13.5574	29	0.001527
GO	RNA modification	GO:0009451	6.65873	18	0.001657
GO	regulation of macromolecule biosynthetic process	GO:0010556	76.4254	109	0.001657
GO	regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	GO:0019219	76.0655	108	0.001925
GO	regulation of biosynthetic process	GO:0009889	78.2251	110	0.00222
GO	regulation of cellular biosynthetic process	GO:0031326	78.2251	110	0.00222
GO	regulation of nitrogen compound metabolic process	GO:0051171	77.0853	108	0.002888
GO	regulation of gene expression	GO:0010468	81.6444	113	0.00327
GO	cell wall organization	GO:0071555	6.65873	17	0.004189
GO	regulation of primary metabolic process	GO:0080090	80.5047	111	0.004189
GO	intrinsic to membrane	GO:0031224	42.8918	66	0.004204
GO	regulation of macromolecule metabolic process	GO:0060255	84.4039	115	0.004788
GO	cell wall modification	GO:0042545	5.03904	14	0.005301

GO	pectinesterase activity	GO:0030599	4.49914	13	0.005542
GO	light-harvesting complex	GO:0030076	1.49971	7	0.005795
GO	heme binding	GO:0020037	10.8579	23	0.005973
GO	transcription regulator activity	GO:0030528	90.4028	121	0.006118
GO	transcription	GO:0006350	79.1849	108	0.006242
GO	chloroplast ribulose bisphosphate carboxylase complex	GO:0009573	0.239954	3	0.008199
GO	ribulose-bisphosphate carboxylase activity	GO:0016984	0.239954	3	0.008199
GO	ribulose bisphosphate carboxylase complex	GO:0048492	0.239954	3	0.008199
GO	regulation of cellular metabolic process	GO:0031323	83.804	110	0.017456
GO	fatty acid (omega-1)-hydroxylase activity	GO:0008393	0.299943	3	0.017654
GO	water transmembrane transporter activity	GO:0005372	1.85965	7	0.017901
GO	water channel activity	GO:0015250	1.85965	7	0.017901
GO	enzyme inhibitor activity	GO:0004857	5.87888	14	0.018255
GO	lipase activity	GO:0016298	5.33898	13	0.020895
GO	response to red light	GO:0010114	2.99943	9	0.022029
GO	response to far red light	GO:0010218	2.45953	8	0.022066
GO	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	GO:0016705	8.69834	18	0.022086
GO	ncRNA processing	GO:0034470	8.09846	17	0.024415
GO	regulation of metabolic process	GO:0019222	90.5228	116	0.025807
GO	regulation of transcription, DNA-dependent	GO:0006355	39.4125	57	0.025948
GO	anther development	GO:0048653	1.07979	5	0.02641
GO	proline dehydrogenase activity	GO:0004657	0.119977	2	0.026629
GO	ureide metabolic process	GO:0010135	0.119977	2	0.026629

GO	ureide catabolic process	GO:0010136	0.119977	2	0.026629
GO	release of seed from dormancy	GO:0048838	0.119977	2	0.026629
GO	exocytic vesicle	GO:0070382	0.119977	2	0.026629
GO	central vacuole	GO:0042807	0.359931	3	0.027437
GO	gibberellin 2-beta-dioxygenase activity	GO:0045543	0.359931	3	0.027437
GO	response to blue light	GO:0009637	3.17939	9	0.028784
GO	regulation of RNA metabolic process	GO:0051252	39.8324	57	0.029511
GO	iron ion binding	GO:0005506	14.2173	25	0.031033
GO	photosynthesis, light harvesting	GO:0009765	1.13978	5	0.031033
GO	DNA binding	GO:0003677	112.179	138	0.040151
GO	ferric-chelate reductase activity	GO:0000293	0.41992	3	0.041963
GO	oxidoreductase activity, oxidizing metal ions	GO:0016722	0.41992	3	0.041963
GO	oxidoreductase activity, oxidizing metal ions, NAD or NADP as acceptor	GO:0016723	0.41992	3	0.041963

Table 4.6 Enrichment Analysis of Genes Upregulated in *ttg2-1* Compared to L*er*.

Category	Subcategory	Subcategory alternative name	expected	observed	p-value
KEGG	Phenylpropanoid biosynthesis	940	7.90009	20	0.00217
KEGG	Drug metabolism - cytochrome P450	982	2.27427	9	0.004384
KEGG	Metabolism of xenobiotics by cytochrome P450	980	2.15457	8	0.013541
GO	defense response	GO:0006952	102.867	176	3.99E-12
GO	endomembrane system	GO:0012505	391.883	493	2.00E-07
GO	innate immune response	GO:0045087	42.4636	79	4.33E-07

GO	response to salicylic acid stimulus	GO:0009751	20.738	47	5.46E-07
GO	immune response	GO:0006955	45.7554	83	6.45E-07
GO	immune system process	GO:0002376	45.92	83	7.57E-07
GO	response to biotic stimulus	GO:0009607	91.017	140	1.86E-06
GO	transmembrane receptor activity	GO:0004888	22.3839	47	6.76E-06
GO	response to other organism	GO:0051707	85.2564	130	8.13E-06
GO	multi-organism process	GO:0051704	110.109	157	3.65E-05
GO	protein amino acid phosphorylation	GO:0006468	130.024	180	4.63E-05
GO	receptor activity	GO:0004872	27.6507	52	6.07E-05
GO	kinase activity	GO:0016301	194.049	253	6.33E-05
GO	programmed cell death	GO:0012501	28.8028	53	9.66E-05
GO	calcium ion binding	GO:0005509	28.9674	53	0.000115
GO	response to stimulus	GO:0050896	509.728	592	0.000281
GO	cell death	GO:0008219	34.0697	58	0.000411
GO	death	GO:0016265	34.0697	58	0.000411
GO	systemic acquired resistance	GO:0009627	6.08974	17	0.000563
GO	post-translational protein modification	GO:0043687	169.855	218	0.000739
GO	response to stress	GO:0006950	293.295	354	0.000866
GO	defense response, incompatible interaction	GO:0009814	17.9401	35	0.000866
GO	phosphorylation	GO:0016310	145.002	189	0.000949
GO	signal transduction	GO:0007165	127.72	169	0.001009
GO	response to chitin	GO:0010200	19.5859	37	0.001093
GO	secondary metabolic process	GO:0019748	58.4286	87	0.001132
GO	transferase activity, transferring phosphorus-containing groups	GO:0016772	223.51	276	0.001202
GO	phosphorus metabolic process	GO:0006793	157.84	202	0.001471
GO	cell recognition	GO:0008037	4.27928	13	0.001471
GO	recognition of pollen	GO:0048544	4.27928	13	0.001471

GO	chitinase activity	GO:0004568	2.79799	10	0.001666
GO	regulation of defense response	GO:0031347	10.369	23	0.001739
GO	phosphate metabolic process	GO:0006796	157.675	201	0.001845
GO	signaling process	GO:0023046	130.353	170	0.001846
GO	signal transmission	GO:0023060	130.353	170	0.001846
GO	pollen-pistil interaction	GO:0009875	4.93763	14	0.001869
GO	response to bacterium	GO:0009617	39.3365	62	0.002043
GO	phenylpropanoid metabolic process	GO:0009698	22.5485	40	0.002258
GO	aromatic compound biosynthetic process	GO:0019438	29.4612	49	0.002488
GO	protein modification process	GO:0006464	198.328	245	0.002582
GO	phenylpropanoid biosynthetic process	GO:0009699	17.6109	33	0.002646
GO	signaling	GO:0023052	153.725	195	0.002784
GO	glutathione transferase activity	GO:0004364	7.07727	17	0.003937
GO	regulation of response to stress	GO:0080134	12.3441	25	0.003974
GO	signal transducer activity	GO:0004871	49.3763	73	0.004142
GO	molecular transducer activity	GO:0060089	49.3763	73	0.004142
GO	cellular aromatic compound metabolic process	GO:0006725	49.2117	72	0.005922
GO	response to fungus	GO:0009620	26.0049	43	0.005952
GO	toxin metabolic process	GO:0009404	6.7481	16	0.00633
GO	toxin catabolic process	GO:0009407	6.7481	16	0.00633
GO	molecular_function	GO:0003674	3204.03	3247	0.012559
GO	apoptosis	GO:0006915	18.763	32	0.015486
GO	macromolecule modification	GO:0043412	220.218	261	0.016106
GO	response to carbohydrate stimulus	GO:0009743	30.6133	47	0.016528
GO	cellular metal ion homeostasis	GO:0006875	5.43139	13	0.016693

GO	metal ion homeostasis	GO:0055065	5.43139	13	0.016693
GO	response to organic substance	GO:0010033	158.498	193	0.01795
GO	cell wall macromolecule catabolic process	GO:0016998	3.12717	9	0.01908
GO	cellular amino acid derivative metabolic process	GO:0006575	38.5135	56	0.022309
GO	transferase activity	GO:0016740	391.883	442	0.022529
GO	sugar binding	GO:0005529	9.71067	19	0.023258
GO	protein tyrosine kinase activity	GO:0004713	32.4238	48	0.027616
GO	protein serine/threonine kinase activity	GO:0004674	99.0818	125	0.028745
GO	cellular calcium ion homeostasis	GO:0006874	2.79799	8	0.030784
GO	calcium ion homeostasis	GO:0055074	2.79799	8	0.030784
GO	ATPase activity, coupled to transmembrane movement of ions, phosphorylative mechanism	GO:0015662	6.58351	14	0.032705
GO	plant-type hypersensitive response	GO:0009626	5.92516	13	0.032963
GO	response to jasmonic acid stimulus	GO:0009753	23.8652	37	0.032996
GO	sulfotransferase activity	GO:0008146	1.81046	6	0.038903
GO	4-coumarate-CoA ligase activity	GO:0016207	1.81046	6	0.038903
GO	DNA catabolic process	GO:0006308	1.3167	5	0.039077
GO	pyrimidine base catabolic process	GO:0006208	0.493763	3	0.039349
GO	uracil catabolic process	GO:0006212	0.493763	3	0.039349
GO	uracil metabolic process	GO:0019860	0.493763	3	0.039349
GO	host programmed cell death induced by symbiont	GO:0034050	6.08974	13	0.039671
GO	carbohydrate binding	GO:0030246	20.4089	32	0.044935
GO	protein kinase activity	GO:0004672	129.201	156	0.04813

GO	defense response to fungus	GO:0050832	19.7505	31	0.048366
GO	intracellular ligand-gated ion channel activity	GO:0005217	2.46882	7	0.049947
GO	amine-lyase activity	GO:0016843	2.46882	7	0.049947
GO	strictosidine synthase activity	GO:0016844	2.46882	7	0.049947

Table 4.7 Relative Expression of Known Trichome Genes.

		GL2 /COL	TTG2 /LER				GL2 /COL	TTG2 /LER
AT3G27920	GL1	0.746	0.634		AT1G22620	SAC1	0.967	1.167
AT1G79840	GL2	0.823	0.634		AT2G03220	MUR2	1.245	1.015
AT5G41315	GL3	2.645	0.906		AT2G02480	STI	1.365	1.163
AT1G63650	EGL3	0.746	1.440		AT1G68020	CPS1	0.984	0.861
AT5G24520	TTG1	0.854	1.012		AT2G31660	SAD2	0.839	1.378
AT2G37260	TTG2	0.622	0.158		AT3G20780	HYP6	1.365	0.725
AT5G40330	MYB23	0.088	0.453		AT5G02820	RHL2	1.073	1.838
AT3G13540	MYB5	0.293	INFINITE		AT5G06700	TBR	0.835	0.823
AT5G53200	TRY	1.764	2.283		AT3G21560	BRT	0.920	1.460
AT2G46410	CPC	0.230	0.317		AT4G26690	SHV3	0.957	0.880
AT1G01380	ETC1	1.175	3.172		AT5G55480	SVL1	1.093	0.575
AT2G30420	ETC2	4.233	1.776		AT4G14960	Lefty1	1.072	1.062
AT4G01060	ETC3	3.527	0.994		AT5G65930	ZWI	0.957	1.468
AT2G30432	TCL1	0.440	0.190		AT1G04820	Lefty2	1.284	0.833
AT1G73360	HDG11	1.175	0.951		AT2G46600	KIC	1.353	2.526
AT1G17920	HDG12	1.411	1.048		AT2G34560	FRA2	0.961	0.614
AT3G01140	NOK	0.355	0.938		AT2G30410	KIS	0.904	1.004
AT4G16340	SPK1	0.988	1.010		AT5G20490	MYO XIK	1.167	0.789
AT5G64930	CPR5	1.190	2.706		AT5G18410	PIR	1.029	0.541
AT5G57800	WAX2	0.993	0.968		AT4G01710	CRK	1.399	1.080
AT5G42080	ADL1	1.060	0.875		AT3G27000	WRM	1.560	2.440
AT5G04470	SIM	1.266	0.749		AT1G13180	DIS1	0.979	2.379
AT3G57860	PYM	1.679	1.203		AT1G30825	DIS2	0.768	1.098
AT4G38600	KAK	1.056	1.188		AT2G35110	GRL	1.018	1.346

AT3G11540	SPY	1.002	0.651		AT2G38440	DIS3	0.924	1.546
AT1G01510	AN	1.117	0.886		AT1G05230	HDG2	0.911	1.434
AT1G4940	SCD1	1.411	1.131		AT1G56580	SVB	0.723	0.720
AT1G15570	CYCA2;3	1.380	0.965		AT1G64690	BLT	0.661	2.537
AT1G03060	SPI	1.228	1.702		AT5G23940	PEL3	0.588	0.570

Table 4.8 T-DNA Lines for Possible GL2 Targets.

CS863534	SALK_061690	SALK_022170
SALK_045206	WiscDsLox508E02	SALK_065561
SALK_015187	SALK_019169	SALK_041400
SALK_085330	CS862629	SALK_030664
SALK_008156	SALK_066589	SALK_144729
CS863372	SALK_012944	SALK_137991
SALK_065223	SAIL_408_D08	SAIL_360_B08
SALK_100085	SALK_041794	SALK_016593
WiscDsLox343E01	SALK_004973	SALK_138985
SALK_041449	SALK_061487	SALK_091664
SAIL_251_H06	SAIL_73_H02	SALK_060495
SALK_146182	SALK_136090	SALK_124913
SALK_007001	SALK_148217	SALK_109929
SAIL_342_B04	WiscDsLox508C12	SALK_084661
SALK_092412	WiscDsLox420A06	SALK_025674
SALK_007317	GT_5_108178	SALK_067086
SALK_113433	SALK_148294	WiscDsLoxHs228_04H
SALK_109440	SALK_142051	SALK_104265
SALK_094986	SALK_125518	SALK_070142
SALK_131710	SALK_048798	SALK_059436
SAIL_657_E10	SALK_139336	SALK_048807
WiscDsLox262E08	SAIL_791_F06	GT_5_46415
SALK_120829	WiscDsLox333H08	SALK_082112
SALK_133277	SALK_095684	WiscDsLoxHs090_01B
SALK_032807	SALK_000867	SALK_149961
SALK_050661	SALK_039486	SALK_025297
SALK_055455	SALK_010482	SALK_112714
WiscDsLox7H11	SALK_108053	SALK_000610

SALK_140660	SALK_129240	SALK_038521
SALK_000610	SALK_020755	SM_3_41248
SALK_092160	SALK_007155	SALK_120084
SALK_019289	SAIL_609_A08	SAIL_504_F02
SAIL_897_G10	GT_3_3026	SALK_124211
SALK_056069		

Table 4.9 T-DNA Lines for Possible TTG2 Targets.			
SALK_143587	SALK_017982C	SALK_082546C	SALK_034225C
SALK_008913C	SALK_052444	CS811308	SALK_094892C
SALK_035098C	CS828672	SALK_042624C	SALK_091452C
SALK_117324	SALK_030815C	SALK_084661C	SALK_122584C
SALK_039285C	SALK_088181C	SALK_135819	SALK_046402
CS849240	CS822755	SALK_134251C	CS870763
SALK_057110C	SALK_049136	SALK_070475C	SALK_088205C
SALK_143606C	SALK_097770C	SALK_043788C	SALK_057129
SALK_020252	SALK_150707	SALK_078642	CS821647
SALK_021619C	SALK_020589C	SALK_056899C	SALK_117524
CS872289	SALK_098602C	SALK_104095C	SALK_114666
SALK_088296C	SALK_038166	SALK_115125C	SALK_020840C
CS174586	SALK_046920C	SALK_107788	SALK_009004C
SALK_147678	SALK_041382C	SALK_069786C	SALK_134310C
SALK_047680C	SALK_017554	SALK_002098C	SALK_043562C
SALK_004699C	SALK_098354C	SALK_143403	SALK_043598C
SALK_144639	SALK_080420	SALK_086910	SALK_065216C
SALK_018864C	CS878069	SALK_020571C	SALK_011414C
SALK_081439C	SALK_024296	SALK_001388C	CS879392
SALK_101062C	SALK_109362C	SALK_142648C	CS879418
SALK_025198C	CS817394	SALK_124227C	CS900015
SALK_051583C	CS811757	SALK_009693C	CS906949
SALK_041697C	SALK_130668C	SALK_014055C	CS909715
CS824120	SALK_055480C	SALK_045606C	SALK_069037C
SALK_150162	SALK_024950C	SALK_088578C	SALK_112341C
SALK_111964C	SALK_086894C	CS835492	SALK_042231C

SALK_098205	CS813702	CS862879	SALK_035875C
SALK_047120C	SALK_136607C	CS863480	SALK_030295
SALK_037588C	SALK_061944C	CS871609	SALK_057388
SALK_137214	SALK_098201	CS872805	CS875314
CS817400	CS808251	CS873050	CS875741
CS849745	SALK_055460C	CS873757	

Table 4.10 Primers for qRT-PCR.

Primer Name	Sequence
AT5G06330RT FWD	TCTCACGACAGCCACTCATC
AT5G06330RT REV	CTGTTGGCTCCGGTAAGAAG
AT1G06100RT FWD	CAAACCAATGACACCTCACG
AT1G06100RT REV	TAAACCGAGAGCCTGGAAGA
AT1G70550RT FWD	AACGGGTTGCTACAATTTGC
AT1G70550RT REV	ACCGACTAATGCACCTGACC
AT5G61340RT FWD	TCTACCTCCGGCTTCTCAAA
AT5G61340RT REV	TGCTGAGGATAGGGAGAGGA
AT1G01600RT FWD	TGGAGCAGAAGATGTCGTTG
AT1G01600RT REV	CCACTGCAACTCCCGTATTT
AT5G65300RT FWD	TCGAAGCTTTAGGTGGAGGA
AT5G65300RT REV	CTCGTTGCCGTGAAGTAACA
AT5G33370RT FWD	TGGGATAGGGCTATGCACTC
AT5G33370RT REV	TGGGATGCATGTATTCAGGA
AT5G15160RT FWD	TCGACAAGCACAAACAAAGC
AT5G15160RT REV	TCGGCTTCCTTGTTCAAGTT
AT5G12420RT FWD	CGCGGAAGAGATAGGTGAAG
AT5G12420RT REV	CTCCCAACGAATGGTGAGAT
AT4G17860RT FWD	GAGATGATCGCCCAAGGTTA
AT4G17860RT REV	CCTGTATATCCGCCGAAGAA
AT3G61840RT FWD	CGCAGACGATAATGAGACGA
AT3G61840RT REV	CCCATGAGAAAGGAATCGAA
AT3G17420RT FWD	TGTGGCTGTCAAAAAGTTGC
AT3G17420RT REV	CTTTGCAGTGCCAACAAGAA
AT2G46140RT FWD	TCTGGACGTACCGGTTAAGG
AT2G46140RT REV	CACCAACAACAGGAATGTCG

AT1G22180RT FWD	TCGCTCGCTGATCACTCTAA
AT1G22180RT REV	AGTAATGGCGGCATCTGAAC
AT2G37390RT FWD	GTTTCAGCTGTGTTGCGAGA
AT2G37390RT REV	CGACCTGCTGGTCTGTATCA
AT3G19590RT FWD	CTCGATTGCTGTTTTCACGA
AT3G19590RT REV	CTCAACACATCGCACTGCTT
AT3G22880RT FWD	CTGGGAAAACCCAATTAGCA
AT3G22880RT REV	GGAAGGTTCCCTCTGTGTCA
At4g05520RT FWD	GGAAAGCCAAAGCTCAACAG
At4g05520RT REV	CTGGGATGTCATACCCCAAC
At3g55660RT FWD	CATCTGAGCTTGCGGTTGTA
At3g55660RT REV	TTTTGTCACCATCGGTTTGA
At5g60880RT FWD	GAAGGAAGGGGATCTTTTGC
At5g60880RT REV	TGGAACCCTAAAGCAACTGG
At4g03100RT FWD	AAGGAAACAGTGTCCCAACG
At4g03100RT REV	AGCCAAACAATGCACATCAA
At3g20150RT FWD	AGACGCCTGTAACCATGTCC
At3g20150RT REV	AACCTGGCATGACCAAGAAC
AT3G18170RT FWD	CGGTGGTGATTCAGGTTCTT
AT3G18170RT REV	CCGATGGATCTCTCACGATT

Table 4.11 Primers for ChIP.

Primer Name	Sequence
TRY prom rev	GgggaccactttgtacaagaaagctgggtgTactattgaagtaagaaaag aaaaatag
TRY prom fwd	ggggacaagtttgtacaaaaaagcaggcttgAcatctttgttgttgttgatat aatc
TRY chip fwd	catcgttgaacttgcatctgcc
EGL3 chip fwd	tgaccacagtgtaactagatgac
EGL3 chip rev	ggtttagtattcttccttgacttc
PLDZ chip fwd	gatctcaaacaaagctatataataag
PLDZ chip rev	atatacaggtgcacgaagaaagc
At1g06100 chip fwd	aagattactttgattgtgtgttcgc
At1g06100 chip rev	tttattagttagtgatgatgatttg
At1g01600 chip fwd	tttttgtgctctatgttctcaac

At1g01600 chip rev	tactggtacctaattttgcagcg
At5g33370 chip fwd	cttcaataattcgatgcacaagc
At5g33370 chip rev	Ggaatggagattgatgattatggg
At5g65300 chip fwd	gaaaaaagaaaagcaaacagctg
At5g65300 chip rev	acaactgatatttgtcatataatg
At3g61840 chip fwd	cctcaagctaactaaactgaatag
At3g61840 chip rev	Atctctttcttgtttctgcttattatg

Table 4.12 Primers for Promoter GUS Constructs.

Primer Name	Sequence
At5g65300 prom fwd	ggggacaagtttgtacaaaaaagcaggcttgGctagtgagtgatcctgagg
At5g65300 prom rev	ggggaccactttgtacaagaaagctgggtgCATttttctatcttcttcgttctc
At3g61840 Prom Fwd	ggggacaagtttgtacaaaaaagcaggcttgCCTCAAACAACAGCTTCAAACC
At3g61840 Prom Rev	ggggaccactttgtacaagaaagctgggtgAtctctttcttgtttctgcttattatg
At3g18180 Prom Fwd	ggggacaagtttgtacaaaaaagcaggcttgTgttatggtttttagtcaatttattg
At3g18180 Prom Rev	ggggaccactttgtacaagaaagctgggtgCtcttgactcttaagttacacaaaac
At4g17860 Prom Fwd	ggggacaagtttgtacaaaaaagcaggcttgAgattattgttttgatcgatcttaatc
At4g17860 Prom Rev	ggggaccactttgtacaagaaagctgggtgGaaccttagaggtcgccagttcg
At5g33370 Prom Fwd	ggggacaagtttgtacaaaaaagcaggcttgTgtttatagatttaaaatggtaatatttg
At5g33370 Prom Rev	ggggaccactttgtacaagaaagctgggtgGgaatggagattgatgattatggg
At5g15160 Prom Fwd	ggggacaagtttgtacaaaaaagcaggcttgTCTGCTAGAAGACATgtttgatttg
At5g15160 Prom Rev	ggggaccactttgtacaagaaagctgggtgAacataaatagtcccaccttttgtg
At1g01600 Prom Fwd	ggggacaagtttgtacaaaaaagcaggcttgCTTGTGTGTGGTCCGAAGAAG
At1g01600 Prom Rev	ggggaccactttgtacaagaaagctgggtgGtagctcttttattattgtttcc

	c
At1g06100 Prom Fwd	ggggacaagtttgtacaaaaaagcaggcttgTgcgtaatcatcactcgctc
At1g06100 Prom Rev	ggggaccactttgtacaagaaagctgggtgCggaccctgtcaaggtatg

Table 4.13 Primers for Overexpressing Possible GL2 Targets.

Primer Name	Sequence
AT3G61840 5' fwd	ggggacaagtttgtacaaaaaagcaggcttgATGGTTAAAACCCAAAAAA AGAACAAG
AT3G61840 3' rev	GgggaccactttgtacaagaaagctgggtgTTAATTTACCCAGAAGTAAC ACCTAG
AT4G17860 5' fwd	ggggacaagtttgtacaaaaaagcaggcttgATGGCTTCGTACTGTTTATT ATACATC
AT4G17860 3' rev	GgggaccactttgtacaagaaagctgggtgTTAGATACCACATTGTCCAC CAG
AT5G33370 5' fwd	GgggacaagtttgtacaaaaaagcaggcttgATGACGAACTCGGTGGCT AAG
AT5G33370 3' rev	GgggaccactttgtacaagaaagctgggtgTTAGGTCATGGAATCAACG GTG
AT5G15160 5' fwd	ggggacaagtttgtacaaaaaagcaggcttgATGTCTTCTAGCAGAAGGT CGAG
AT5G15160 3' rev	GgggaccactttgtacaagaaagctgggtgTTATCCATTAATCAAGCTCCT AATAAC
AT1G01600 5' fwd	ggggacaagtttgtacaaaaaagcaggcttgATGGAAATATCCAATGCCA TG
AT1G01600 3' rev	GgggaccactttgtacaagaaagctgggtgTTAAACCACTGCAACTCCCG
AT1G06100 5' fwd	ggggacaagtttgtacaaaaaagcaggcttgATGAGTGAGACCACTAAG GACG
AT1G06100 3' rev	GgggaccactttgtacaagaaagctgggtgTTAACGACGGATAGCCATCT TG
AT3G18180 5' fwd	ggggacaagtttgtacaaaaaagcaggcttgATGACAAAGAAGGATATTC TTTACGATAC
AT3G18180 3' rev	GgggaccactttgtacaagaaagctgggtgTTACACTGACTGATTATGCA ATAGCTC
AT5g65300 5' fwd	ggggacaagtttgtacaaaaaagcaggcttgATGGAATGCAGAAAACACA ACCAC
AT5g65300 3' rev	ggggaccactttgtacaagaaagctgggtgTTAATAAACTCGTTGCCGTG AAG

Chapter 5: GL2's Transcriptional Target Regulates Trichome Accessory Cell Development

Trichome accessory cells or socket cells are a ring of cells that surround the base of a trichome in the epidermis. They are of unique morphology from the rest of the epidermis, however these cells have little known about them. There are on average 8-10 accessory cells around the trichome and these cells differentiate during the last stages of trichome development (Hulskamp and Misera, 1994). Larkin (1996) showed that they are not derived from any set of specific cell divisions, like stomatal guard cells (Nadeau and Sack, 2002), but rather they appear to respond to positional signals that recruit cells adjacent to a trichome to become accessory cells (Larkin et al. 1996). These cells have the appearance of functioning in the structural support of the trichome. In some species the ring of accessory cells will develop asymmetrically in order to shift or tilt the orientation of the trichome (Uphof, 1962), though in *Arabidopsis* accessory cells are uniform in shape.

One of possible targets of GL2, an unknown gene called At5g65300, has shown an effect upon trichome accessory development when overexpressed. This gene was identified through a bioinformatic screen of available expression databases for possible targets of GL2, detailed in chapter 4. At3g65300 is a small unknown protein consisting of 150 amino acids with no known domains. It has no previously published or predicted function and no homologs in *Arabidopsis.* Blast searches reveal that it is present in other

species such as: *Oryza sativa, Zea mays, Nicotiana tabacum,* and *Medicago truncatula*. Based on computer predictions its subcellular localization is unclear, though it is likely localized in the nucleus as two prediction programs point to this result (WoLFPSORT and SubLoc). Other prediction programs (iPSORT, MitoPred, MultiLoc and TargetP) suggest plastid or mitochondria localization which seems less likely.

RESULTS

Overexpression of At5g65300

When At5g65300 is overexpressed using 35S promoter in Col wild type a unique phenotype is produced. Trichomes reside at the top of a pillar of cells which can be taller than the trichome itself (Fig 5.1). These cells have morphology and position suggestive of being trichome accessory cells. This pillar of cells varies in height from one slightly elongated layer of cells to over four large tiers of cells (Fig 5.2). Just as the trichome number increases on the later leaves, the height of the pillar tends to increase from early to later leaves. For example on the first true leaf the pillar is barely noticeable as a slight swelling of the base while on the third true leaf the pillar averages a little over three tiers of accessory cells. While there are great changes to the accessory cells in these transgenic plants the trichomes appear to be unchanged in any way.

The pillars of cells were observed to have a green interior. Sectioning of 35S::At5g65300 leaves was performed to observe the anatomy of the pillar. Longitudinal sections reveal that the accessory cell pillars consist of two or more mesophyll cells between epidermal layers of accessory cells (Fig 5.3). It appears that the palisade layer follows the epidermis up into the pillar with spongy mesophyll filling in the space created

by the altered interior morphology (Fig 5.3). In wild type leaves trichomes generally sit in the plane of the epidermis on top of the palisade layer with accessory cells surrounding and cupping the trichome (Fig 5.3). It appears that the pillar phenotype does not disrupt this general plan as the trichome remains surrounded/cupped by accessory cells while still sitting upon a photosynthetic palisade layer (Fig 5.3).

Overexpression of At5g65300 in *gl2* mutant reveals some interesting aspects of at5g65300's function. In *gl2* mutants the early trichomes do not have accessory cells (Fig 5.4A), which is likely due to the fact that these trichomes never fully mature. The overexpression has no effect on these trichomes. There is no pillar of cells produced likely due to the lack of accessory cells (Fig 5.4B). However as *gl2* mutants age, on later leaves some more developed trichomes appear. These will often branch once, papillate and produce accessory cells (Fig 5.4C). With At5g65300 overexpression these more developed trichomes will produce the pillar of cells (Fig 5.4D).

In wild type plants trichomes develop on the stems usually without branching or accessory cells. There is a slight increase in height of the cells around the trichome yet there is no clear accessory cell morphology displayed surrounding wild type cauline trichomes (Fig 5.5A). When At5g65300 is overexpressed, trichomes on the stem form the same sort of pillar of accessory cells as observed around leaf trichomes (Fig 5.5B). This means that At5g65300 is capable of inducing accessory cell production in the stem where none were previously present. However, the stem trichome pillars appear to be less ordered than the leaf accessory cell pillars. The leaf trichome accessory cell pillars are highly ordered in vertical files of cells with very little disruption within the file (Fig 5.2D). The stem trichome accessory cell pillars are also generally organized into vertical

files of cells near the base of the trichome, but at the top of the pillar they often lose their organization causing the files to shift from the vertical and cells are smaller with altered shape compared both to leaf pillars and the base of the stem pillars (Fig 5.5B). This disorganization appears as if division is still taking place after the maximum height has been reached resulting in the cells near the top stacking up resulting in a shift of file position, cell size and shape in order to accommodate all of the new cells. In this "pile up" theory it would suggest the pillar grows from the base of the trichome and not the base of the pillar as the new cells would pile up near the growing end of a column. Relative elongation of the cells within a file supports a trichome base based growth of the pillar in both stem and leaf based pillars. The cells at the base of a pillar are the longest while the cells at the top, near the base of the trichome are the shortest and roundest.

Localization of At5g65300

At5g65300 promoter GUS shows that it is expressed exclusively in trichomes in shoots. Closer examination of its expression reveals that it is not expressed in trichome accessory cells (Fig 4.4B). Its expression occurs after GL2 is expressed based upon analysis of promoter GUS lines for each. At5g65300 is expressed from at least stage 4 trichomes and like GL2 does not cease expression in trichomes. The timing of this expression places At5g65300 at the correct time to be involved in signaling of accessory cell fate.

While At5g65300's timing is sufficient its position is questionable since it is not expressed in accessory cells yet that is where it appears to function. I hypothesized that it moves from the trichome to the surrounding cells to aid in the development and

specification of the accessory cell fate. To test this hypothesis I introduced YFP tagged At5g65300 construct using microprojectile bombardment and examined whether it moves to the cells around the transformed cell. Previous work (chapter 2) has shown that there is a very low probability of producing clusters of transformed cells using this method. Figure 5.6 shows that At5g65300 is able to move to adjacent cells while being mostly localized in the nucleus of the cell.

Accessory cell developmental timing

Trichome development flows on a set timeline with defined developmental stages first described by Szymanski et al (1998). After the trichome cell fate is decided it takes about 72 hours and 6 stages for the trichome to complete its development. After a cell is recuited to the trichome cell fate it enters stage one where trichomes first undergo endoreduplication and start to grow within the plane of the leaf. In stage two the trichome begins to elongate vertically (Fig 5.7A). Then in stage three the trichome undergoes branching with each branching event occurring separately in time (Fig 5.7B, C and D). At this point the trichome looks like a miniature trichome yet there is no sign of accessory cells. The final stages of trichome development consist of more growth and elongation (Fig 5.7E and F), followed by cell wall thickening and production of papillae on the surface of the trichome in stage six (Fig 5.7G-J). As these final aspects of trichome maturation occurs the first signs of accessory cells are seen sometime between stage five and six.

The cells around the trichome appear the same as the rest of the leaf surface until very late in trichome development. Generally accessory cells make their first appearance

during the production of papillae on the trichome surface (Fig 5.7G). The first definitive sign of accessory cells is the outer surface of the cells around the trichome extending out of the plane of the leaf. This extension of early accessory cells out of the plane of the rest of the leaf epidermis could be occurring through a altered cell shape or a cell division producing a pavement cell and the accessory cell. The cells also change shape to appear block-like. At5g65300 overexpression does not affect the timing of accessory cell development.

During trichome development the nucleus moves up from the base of the trichome to roughly the center of the trichome just below the first branch point. Accessory cell nuclei also appear to move within the cells to a specific position. The nucleus in an accessory cell positions itself as close to the trichome as possible (Fig 5.8). This suggests increased levels of communication between accessory cell and trichome.

Accessory Cell Mutants

While there are no published accessory cell mutant phenotypes, examination of several known trichome mutants for defects in accessory cell formation, either experimentally or by careful examination of figures in the literature, allowed identification of a few mutants with accessory cell phenotypes. As previously noted, *gl2* mutants do not have accessory cells on many of their trichomes. *TTG2* mutants also have some trichomes that appear to never reach maturity that have no accessory cells (Fig 5.9C). A literature search has also identified *gl3-sst* mutants as having no accessory cells (Esch et al. 2003). In all three of these mutants the trichomes without accessory cells also don't have papillae while the trichomes in these mutants with papillae also have

accessory cells. There are several glassy trichome mutants (lacking papillae) however pictures of their accessory cells are not available at this time. Since accessory cells in wild type appear during the production of papillae it is possible these two characteristics of trichome maturation are linked in some way.

SMALLER TRICHOME with VARIBLE BRANCHES (SVB) is a trichome mutant with variable branching that based upon examination of published SEM images, have what appears to be elongated accessory cells (Marks et al. 2009). Based upon published SEM images, MYB103 overexpression seems to also result in slightly elongated accessory cells (Higginson et al. 2003).

KAKTUS (KAK) mutants produce larger trichomes with increased branching and also produce two layers of accessory cells on the first pair of true leaves (Fig 5.9A). The layer further from the trichome is smaller than normal and doesn't always completely surround the trichome. At some positions a single larger accessory cell spans both layers.

During an enhancer/suppressor screen of *ttg2-1* EMS mutagenized plants, two mutants with elongated accessory cells, increased clustering, and abnormal branching were identified, M3-7 and M3-11 (Fig 5.9B). These are still unknown mutants and complementation tests have not been performed.

There are some trichome mutants, like *STICHEL (STI),* that have accessory cells that are wide and flat but otherwise appear normal (Fig 5.9D). Since wild type accessory cells are much more vertically orientated this appears to be a mutant phenotype.

MATERIALS AND METHODS

Overexpression of At5g65300

At5g65300 cDNA from Col was amplified using primers At5g65300 5' Fwd and At5g65300 3' Rev, cloned, sequenced and was recombined into pB7WG2 using Gateway. This construct was transformed into Col wild type, *gl2-1*, and *gl2-5* plants using Agrobacteria floral dip transformation protocol.

Microparticle Bombardment

Microparticle bombardment was performed as previously described in chapter 2 using At5g65300 Col cDNA in pB7WGY2.

SEM

Samples for scanning electron microscopy (SEM) experiments were prepared and visualized as previously described (Payne et al., 2000). Critical point dried specimens were coated with platinum palladium in a Cressington 208 sputter coater and then visualized with a Zeiss Supra 40VP SEM.

DISCUSSION

Work in the previous chapter identified At5g65300 as a direct target of GL2. While At5g65300 T-DNA insertion lines have no phenotype plants overexpressing the protein show a severe accessory cell phenotype. This is consistent with part of the *gl2* mutant phenotypic syndrome because the less developed types of gl2 mutant trichomes do not induce accessory cell differentiation. As a result we hypothesize that GL2 controls the development of the trichome accessory cells through At5g65300 and

probably several other genes. Since At5g65300 is transcribed in the trichome, I hypothesize that the At5g65300 protein moves to the surrounding cells where it induces the accessory cell fate. At5g65300 being most important in the elongation and possible division processes. Since At5g65300 overexpression does not induce cell growth or division in any cells other than trichome accessory cells in wild type plants, it is probable that there is something about trichome accessory cells that make them competent to facilitate At5g65300's function. As a result there has to be some other signal in these cells that makes them competent. I predict that at least one other signal is also coming from trichome that moves to the adjacent cells to mark them as accessory cells and therefore target cells of At5g65300.

During the examination of *gl2-1* mutant trichomes with and without accessory cells, the positions of the cells around the trichome seem to be altered in relation to the plane of the epidermis. In the early trichomes which do not produce accessory cells, the cells around the trichome are level with the rest of the pavement cells. Meanwhile the accessory cells around the more developed trichomes exist above the plane of the pavement cells (Fig 5.10). Whether this is due to the shape change from pavement to accessory cell or a cell division is unclear. What is clear is if this vertical expansion was amplified it would phenocopy the overexpression phenotype of At5g65300 suggesting that this is At5g65300's normal function.

Overexpression of At5g65300 in *gl2* mutants does not result in accessory cells surrounding the trichomes that do not normally induce them. However overexpression does induce accessory cells in wild type on the stem where they are also not normally present. This suggests that the stem trichomes are much more susceptible than *gl2*

mutant trichomes to produce accessory cells. It is unclear what is different between these two sets of trichomes.

One thing that seems clear about accessory cell development is that it consists of both trichome and leaf epidermis based genes. This is illustrated by the fact that trichomes produced by MYB23 overexpression on the cotyledons do not produce accessory cells (Kirik et al. 2001). They also don't form in the subepidermal trichomes induced by GL1 overexpression in *TRIPTYCHON (TRY)* mutants (Schnittger et al. 1998). Finally stem trichomes don't normally produce accessory cells either. While At5g65300, GL2, TTG2 and GL3 are the only known genes involved in accessory cell development all are produced in the trichome, which would not explain why trichomes produced in locations other than the leaf epidermis do not produce accessory cells.

The epidermis has long been studied to gain insight into the genetic and molecular mechanisms leading to the differentiation and patterning of cells. However one of the cell types in the epidermis has long been ignored. Accessory cells have long been observed but no molecular or genetic analysis has been performed other than Larkin's study separating accessory lineage from trichome cell lineage (1996). The work presented here illustrates the probability of an interesting and complicated network of regulation involved in the development of the accessory cell fate opening up a new unexplored field of epidermal cell fate.

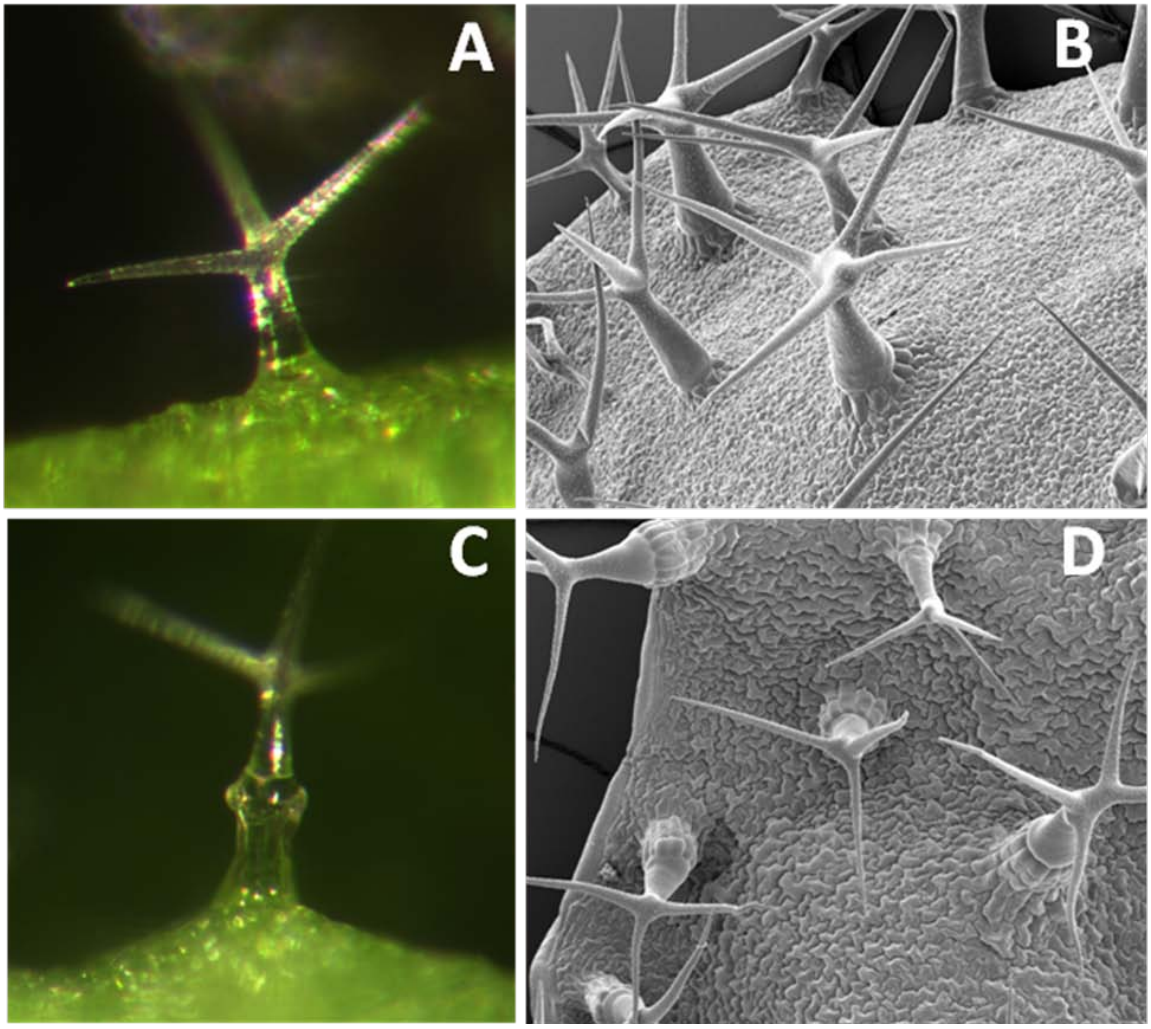

Fig 5.1. Overexpression of At5g65300.

Trichomes and trichome accessory cells of A-B.) wild type and C-D.) At5g65300 overexpression lines shown in either dissecting scope or SEM images. Note the pillar of accessory cells holding up the trichomes of the At5g65300 overexpression plants.

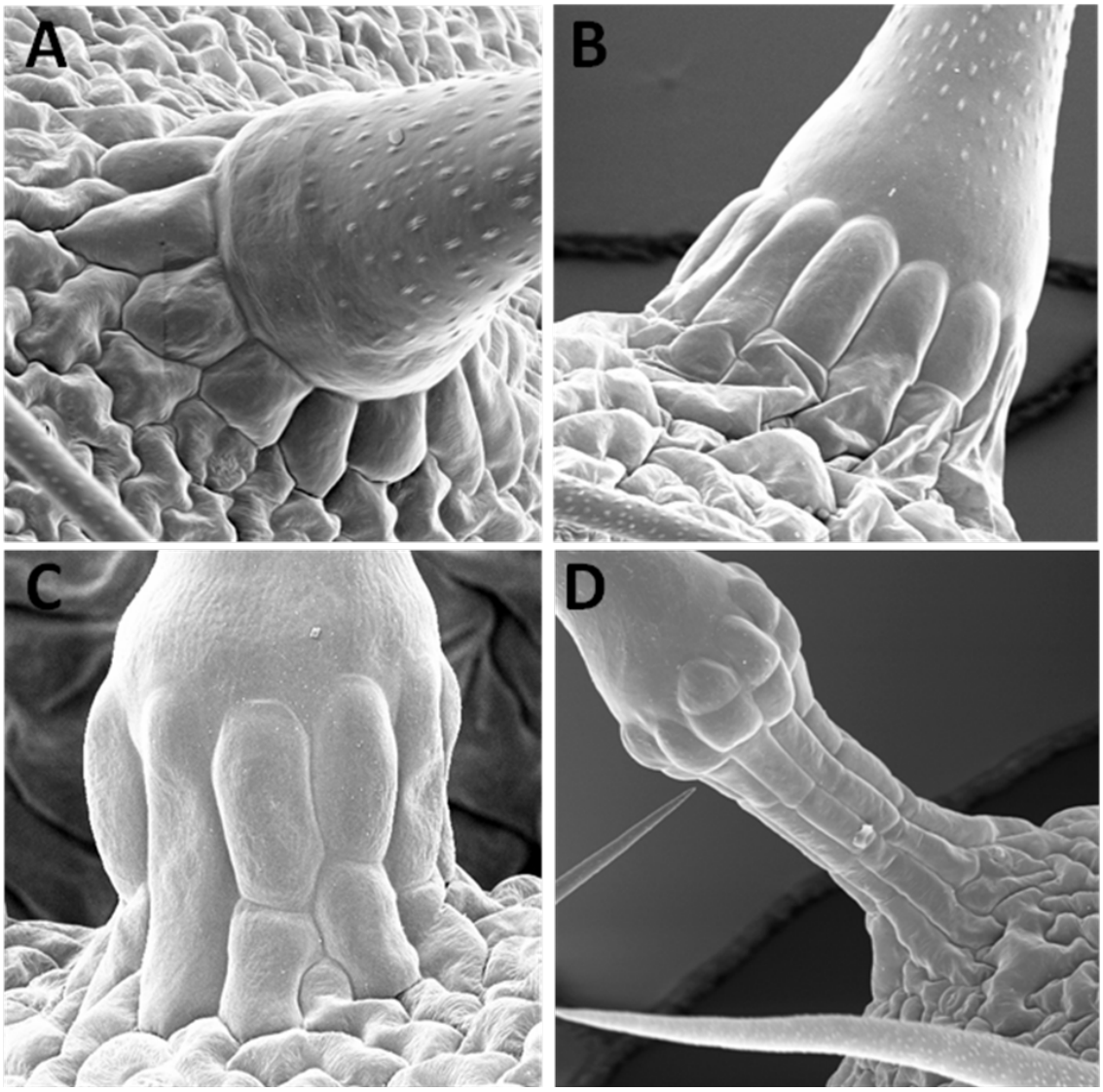

Fig 5.2. Variable Height of the Accessory Cell Pillars.

A.) SEM of trichome accessory cells in wild type Col plants. B-D) SEM of trichome accessory cells in plants overexpressing At3g65300 in Col. Phenotype changes from 2nd true leaves (B.) to 3rd (C.) and 4th (D.) leaves in primary transgenic plants.

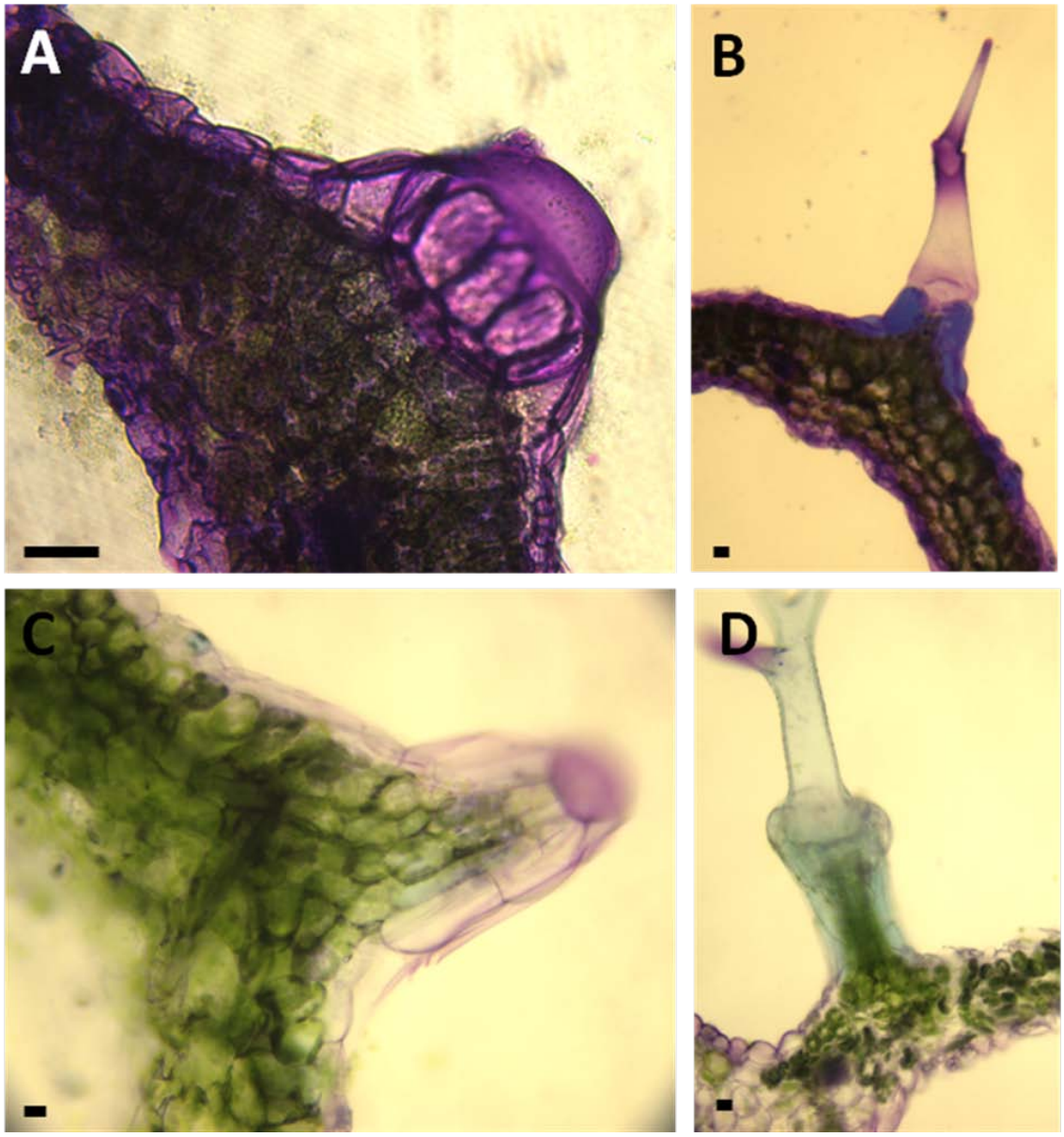

Fig 5.3. Longitudinal Sections of the Pillars.

Touline Blue stained 50 -60 um sections of Col wild type A.) and Col plants overexpressing At5g65300 B-D). Note how the trichome remains attached to the mesophyll layer no matter how tall the pillar gets. Size bars represent 20 um.

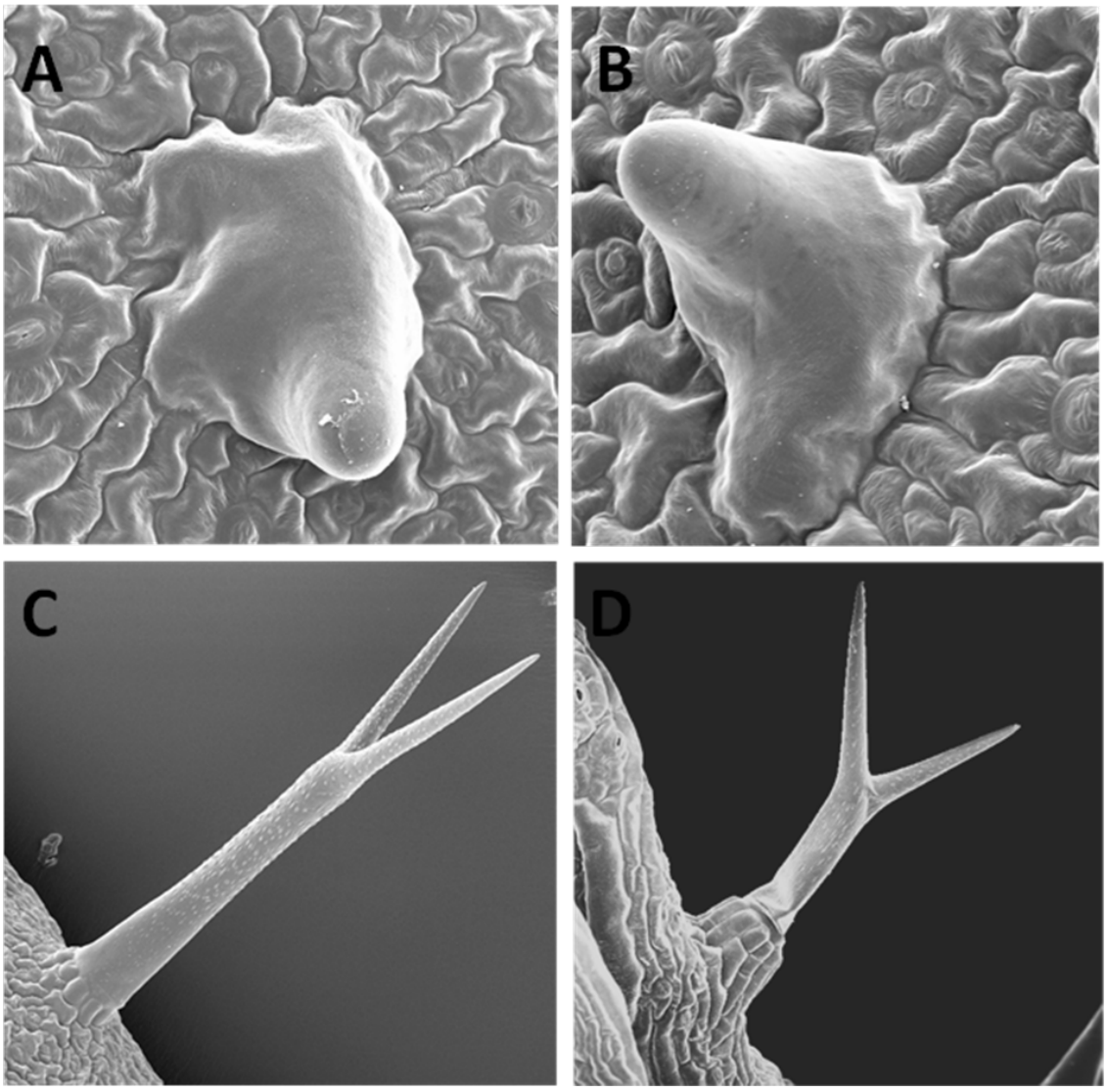

Fig 5.4. Overexpression of At5g65300 in *gl2-1*

A.) SEM of *gl2-1* nub-like trichomes which lack accessory cells. B.) *gl2-1* nub trichomes overexpressing At3g65300 which still produce no accessory cells. C.) SEM of *gl2-1* more developed trichomes which produce accessory cells. D.) *gl2-1* more developed trichomes overexpressing At3g65300 which produce a pillar of accessory cells.

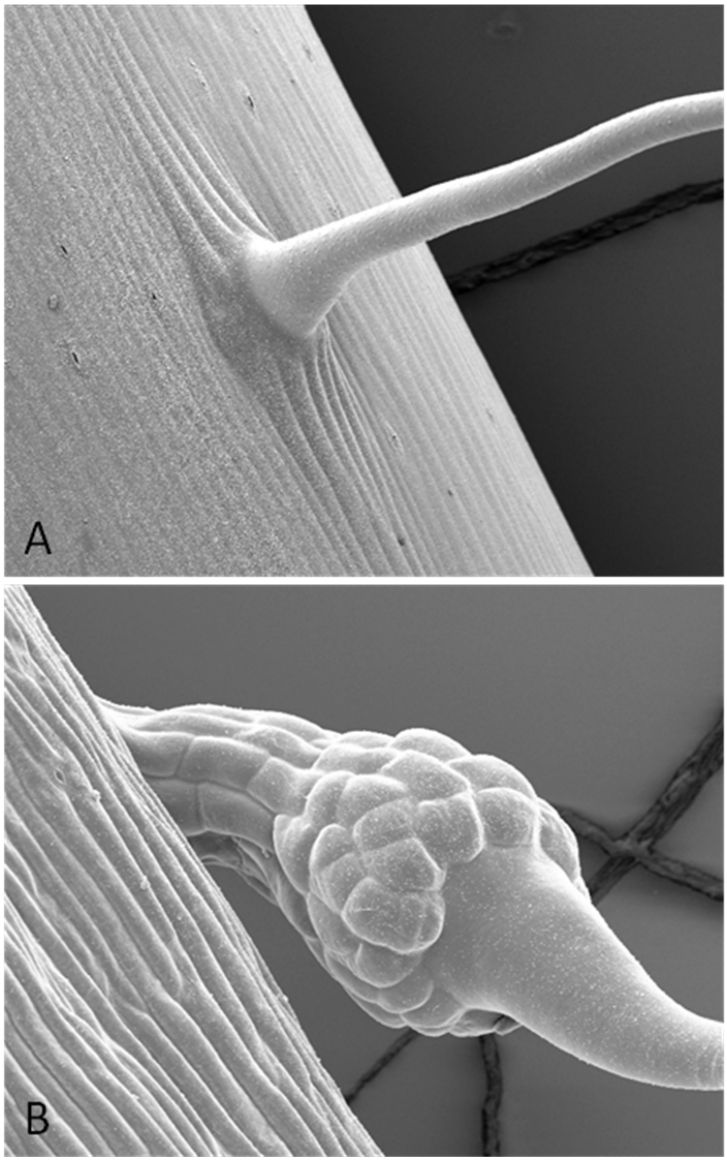

Fig 5.5. Stem Trichomes.

SEM of Stem Trichome Accessory Cells. In A.) Col wild type there are no clear accessory cell morphology observed, however in B.) 35S::At5g65300 in Col wild type the pillar of accessory cells is formed but with more disorder than in the pillars of leaf trichomes near the base of the trichome.

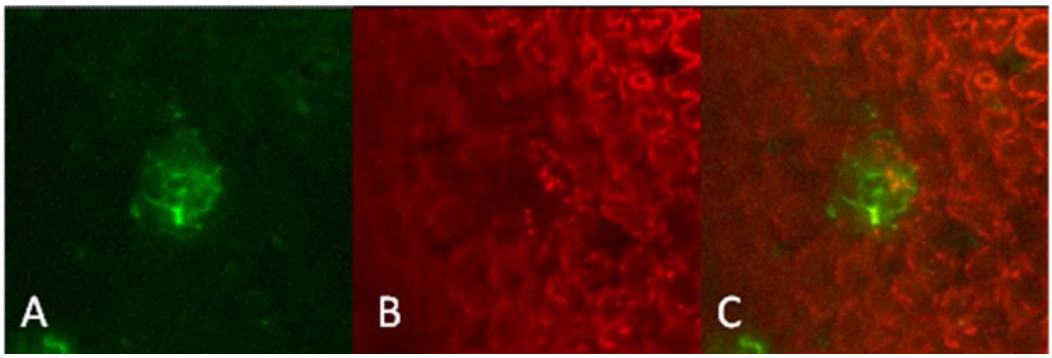

Fig 5.6. Intracellular Trafficking of At5g65300 in Leaf Epidermis.

Confocal images of microparticle bombardment with 35S::YFP-At5g65300 into the leaf epidermis. YFP shown in green and chlorophyll autofluoresence shown in red. A.) YFP Fluoresence. B.) chlorophyll fluoresence. C.) both merged.

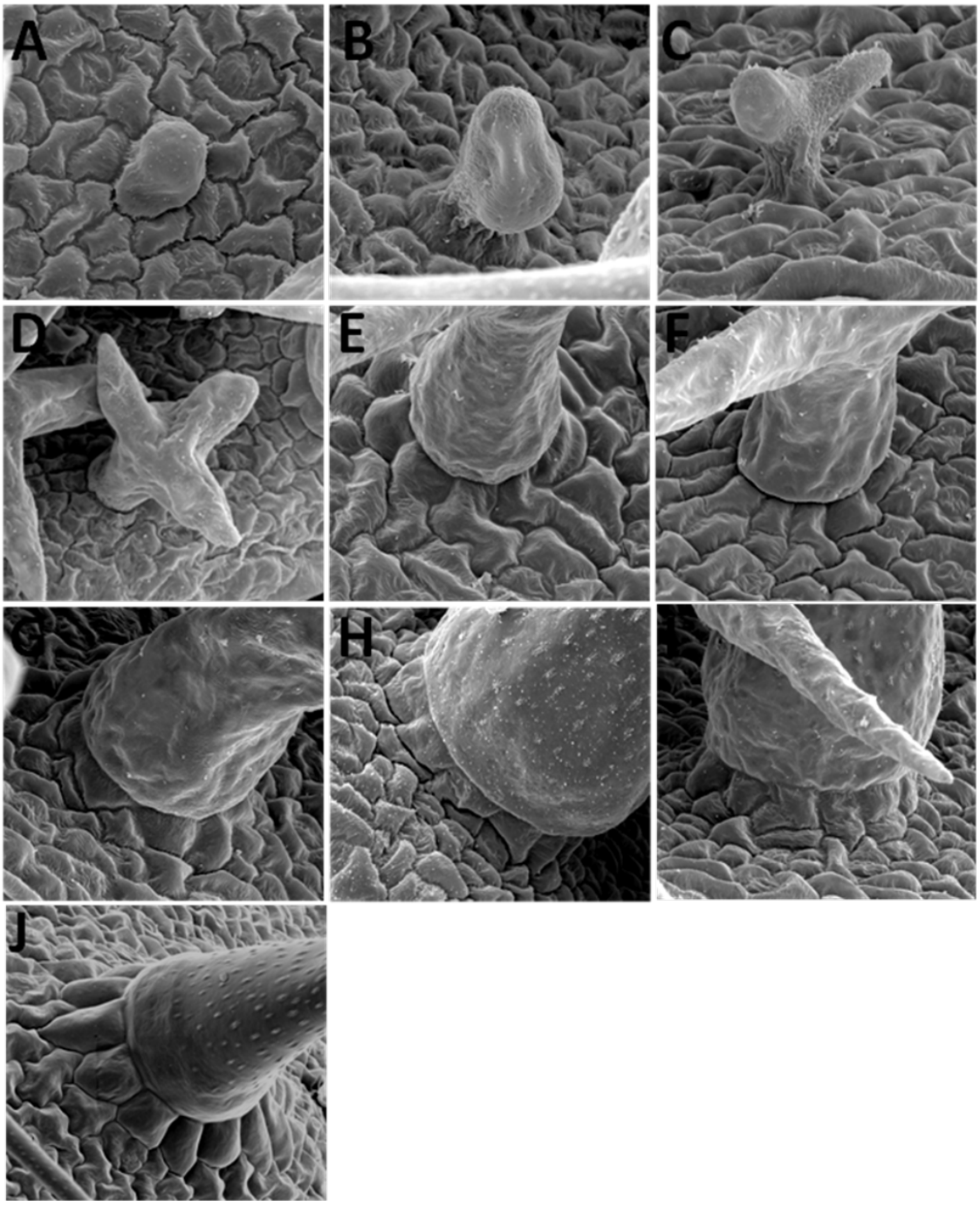

Fig 5.7. Timing of Trichome and Accessory Cell Development.

SEM of Columbia Wild Type Trichome Accessory Cell Development. A-J.) Sequential pictures showing trichome development from trichome initial through maturity. Cells around fig F may show evidence of accessory cells just starting to develop but aren't clear until in figure G.

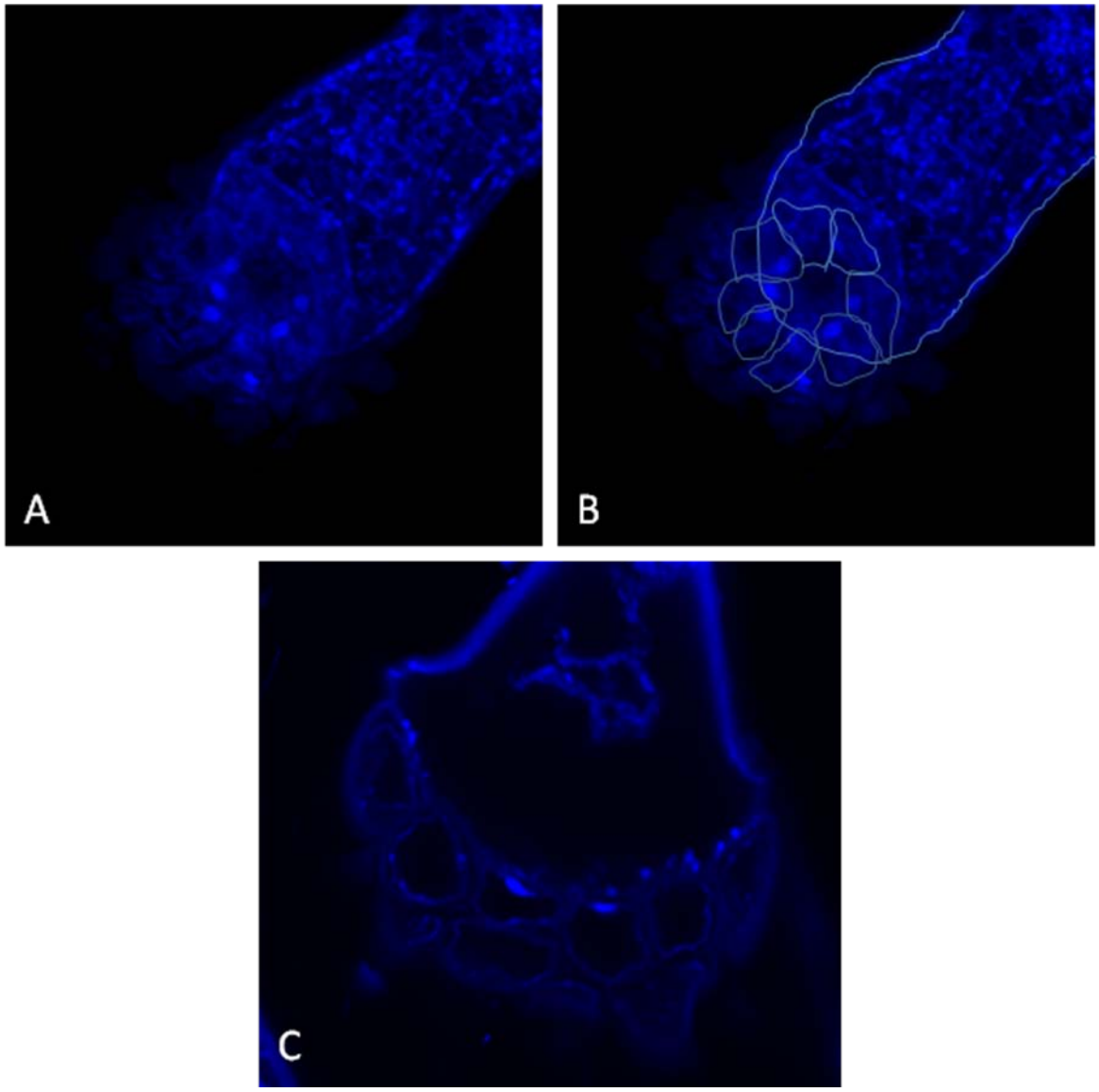

Fig 5.8. Localization of the Nucleus in Accessory Cells.

Confocal images of DAPI stained Col wild type trichomes. The nucleus, stained blue, in the accessory cells is directed to a position adjacent to the trichome. A.) and B.) are the same image of a stage 5 trichome but in B.) the trichome and the surrounding accessory cells have been outlined in white. C.) A fully mature stage 6 trichome.

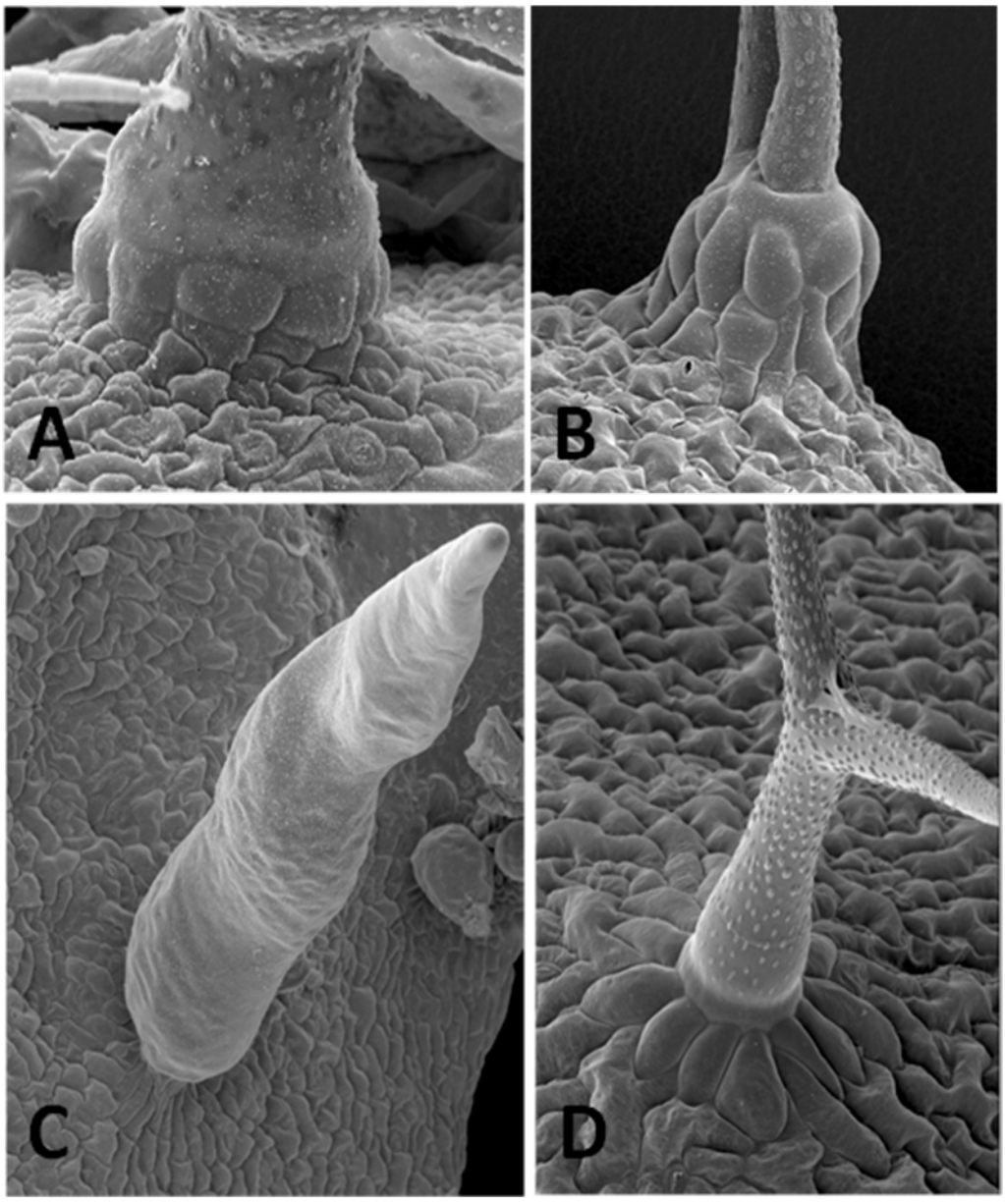

Fig 5.9. Mutants With Accessory Cell phenotypes.

SEM of trichome mutants With abnormal accessory cells. A.) kak (SALK_037636), B.) M3-7, C.) ttg2-1 and D.) sti mutant trichomes.

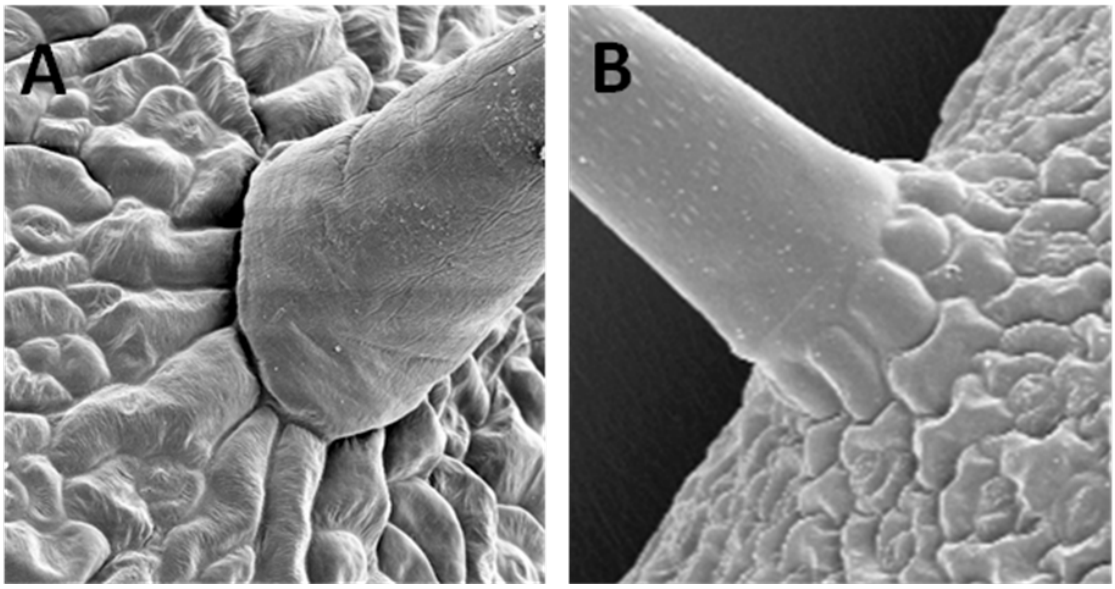

Fig 5.10. Accessory cells of *gl2-5*

A. SEM of *gl2-5* nub-like trichomes which lack accessory cells. B. SEM of *gl2-5* more developed trichomes which produce accessory cells. Accessory cells which surround the trichome are extended out of the plane of the leaf surface while cells surrounding trichomes which don't become accessory cells remain in the plane of the leave with the rest of the pavement cells.

References

Abe, M., Takahashi, T. and Komeda, Y. (2001) Identification of a cis-regulatory element for L1 layer-specific gene expression, which is targeted by an L1-specific homeodomain protein. *The Plant Journal* **26**, 487-494.

Alpy, F. and Tomasetto, C. (2005) Give lipids a START: the StAR-related lipid transfer (START) domain in mammals. *Journal of Cell Science* **118**, 2791-2801.

Baudry, A., Heim, M. A., Dubreucq, B., Caboche, M., Weisshaar, B. and Lepiniec, L. (2004). TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in Arabidopsis thaliana. *Plant J* **39**, 366-80.

Baudry, A., Caboche, M. and Lepiniec, L. (2006). TT8 controls its own expression in a feedback regulation involving TTG1 and homologous MYB and bHLH factors, allowing a strong and cell-specific accumulation of flavonoids in Arabidopsis thaliana. *Plant J*. **46**, 768-79.

Bernhardt, C., Zhao, M., Gonzalez, A., Lloyd, A. and Schiefelbein, J. (2005). The bHLH genes GL3 and EGL3 participate in an intercellular regulatory circuit that controls cell patterning in the Arabidopsis root epidermis. *Development* **132**, 291-8.

Caro, E., Castellano, M. M. and Gutierrez, C. (2007). A chromatin link that couples cell division to root epidermis patterning in Arabidopsis. *Nature* **447**, 213-7.

Chandler, J., Cole, M., Flier, A., Grewe, B., and Werr, W. (2007) The AP2 transcription factors DORNROSCHEN and DORNROSCHEN-LIKE redundantly control Arabidopsis embryo patterning via interaction with PHAVOLUTA. *Development* **134**, 1653-1662

Costa, S., and Dolan, L. (2003) Epidermal patterning genes are active during embryogenesis in Arabidopsis. *Development* **130**, 2893–2901.

de Vetten, N., Quattrocchio, F., Mol, J. and Koes, R. (1997). The an11 locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals. *Genes Dev* **11**, 1422-34.

Debeaujon, I., Nesi, N., Perez, P., Devic, M., Grandjean, O., Caboche, M., and Lepiniec, L. (2003) Proanthocyanidin-Accumulating Cells in Arabidopsis Testa: Regulation of Differentiation and Role in Seed Development. *Plant Cell* **15**, 2514-2531.

Di Cristina, M., Sessa, G., Dolan, L., Linstead, P., Baima, S., Ruberti, I. and Morelli, G. (1996). The Arabidopsis Athb-10 (GLABRA2) is an HD-Zip protein required for regulation of root hair development. *Plant J* **10**, 393-402.

Esch, J.J., Chen, M., Sanders, M., Hillestad, M., Ndkium, S., Idelkope, B., Neizer, J., and Marks, M.D. (2003) A contradictory GLABRA3 allele helps define gene interactions controlling trichome development in Arabidopsis. *Development* **130**, 5885-5894.

Eulgem, T., Rushton, P.J., Robatzek, S., and Somssich, I.E. (2000) The WRKY superfamily of plant transcription factors. *Trends in Plant Science* **5**, 199-206.

Garcia, D., Fitz Gerald J.N., and Berger, F. (2005) Maternal Control of Integument Cell Elongation and Zygotic Control of Endosperm Growth Are Coordinated to Determine Seed Size in Arabidopsis. *Plant Cell* **17**, 52-60.

Gonzalez, A., Zhao, M., Leavitt, J.M., Lloyd A.M. (2008) Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/MYB transcriptional complex in Arabidopsis seedlings. *Plant Journal* **53**, 814-827

Gonzalez, A., Mendenhall, J., Huo, Y. and Lloyd, A. (2009) TTG1 complex MYBs, MYB5 and TT2, control outer seed coat differentiation. *Developmental Biology* **325** (2), 412-421.

Higginson, T., Li, F.S., and Parish, R.W. (2003) AtMYB103 regulates tapetum and trichome development in Arabidopsis thaliana. *Plant Journal* **35**, 177-192.

Hulskamp, M., and Misera, S. (1994) Genetic Dissection of Trichome Cell Development in Arabidopsis. *Cell* **76**, 555-566.

Hulskamp, M. and Schnittger, A. (1998). Spatial regulation of trichome formation in Arabidopsis thaliana. *Semin Cell Dev Biol* **9**, 213-20.

Hulskamp, M., Schnittger, A. and Folkers, U. (1999). Pattern formation and cell differentiation: trichomes in Arabidopsis as a genetic model system. *Int Rev Cytol* **186**, 147-78.

Ishida, T., Hattori, S., Sano, R., Inoue, K., Shirano, Y., Hayashi, H., Shibata, D., Sato, S., Kato, T., Tabata, S., Okada, K., and Wada, T. (2007) Arabidopsis TRANSPARENT TESTA GLABRA2 is directly regulated by R2R3 MYB transcription factors and is involved in

regulation of GLABRA2 transcription in epidermal differentiation. *Plant Cell* **19**, 2531-2543

Jakoby, M.J., Falkenhan, D., Mader, M.T., Brininstool, G., Wischnitzki, E., Platz, N., Hudson, A., Hülskamp, M., Larkin, J., and Schnittger, A. (2008) Transcriptional Profiling of Mature Arabidopsis Trichomes Reveals That NOECK Encodes the MIXTA-Like Transcriptional Regulator MYB106. *Plant Physiology* **148**, 1583-1602.

Johnson, C.S., Kolevski, B., and Smyth, D.R. (2002) TRANSPARENT TESTA GLABRA2, a Trichome and Seed Coat Development Gene of Arabidopsis, Encodes a WRKY Transcription Factor. *Plant Cell* **14**, 1359-1375.

Karimi, M., Inze, D., and Depicker A. (2002) GATEWAY((TM)) vectors for Agrobacterium-mediated plant transformation. *Trends Plant Sci.* **7**, 193**–**195.

Kim, J., Harter, K., and Theologis, A. (1997) Protein-protein interactions among Aux/IAA proteins. *PNAS* **94**, 11786-11797.

Kirik, V., Schnittger, A., Radchuk, V., Adler, K., Hulskamp, M., and Baumlein, H. (2001) Ectopic Expression of the Arabidopsis AtMYB23 Gene Induces Differentiation of Trichome Cells. *Developmental Biology* **235**, 366-377.

Kirik, V., Lee, M.M., Wester, K., Herrmann, U., Zheng, Z., Oppenheimer, D., Schiefelbein, J., and Hulskamp, M. (2005) Functional diversification of MYB23 and GL1 genes in trichome morphogenesis and initiation. *Development* **132**, 1477-1485.

Kirik, V., Simon, M., Hulskamp, M. and Schiefelbein, J. (2004a). The ENHANCER OF TRY AND CPC1 gene acts redundantly with TRIPTYCHON and CAPRICE in trichome and root hair cell patterning in Arabidopsis. *Dev Biol* **268**, 506-13.

Kirik, V., Simon, M., Wester, K., Schiefelbein, J. and Hulskamp, M. (2004b). ENHANCER of TRY and CPC 2 (ETC2) reveals redundancy in the region-specific control of trichome development of Arabidopsis. *Plant Mol Biol* **55**, 389-98.

Koshino-Kimura, Y., Wada, T., Tachibana, T., Tsugeki, R., Ishiguro, S. and Okada, K. (2005). Regulation of CAPRICE transcription by MYB proteins for root epidermis differentiation in Arabidopsis. *Plant Cell Physiol* **46**, 817-26.

Larkin, J.C., Young, N., Prigge, M., and Marks, M.D. (1996) The control of trichome spacing and number in Arabidopsis. *Development* **122**, 997-1005.

Larkin, J. C., Marks, M. D., Nadeau, J. and Sack, F. (1997). Epidermal cell fate and patterning in leaves. *Plant Cell* **9**, 1109-20.

Larkin, J. C., Brown, M. L. and Schiefelbein, J. (2003). How do cells know what they want to be when they grow up? Lessons from epidermal patterning in Arabidopsis. *Annu Rev Plant Biol* **54**, 403-30.

Lee, M. M. and Schiefelbein, J. (1999) WEREWOLF, a MYB-related protein in Arabidopsis, is a position-dependent regulator of epidermal cell patterning. *Cell* 99, 473–483.

Lee, M. M. and Schiefelbein, J. (2001). Developmentally distinct MYB genes encode functionally equivalent proteins in Arabidopsis. *Development* **128**, 1539-46.

Lee, M. M. and Schiefelbein, J. (2002). Cell pattern in the Arabidopsis root epidermis determined by lateral inhibition with feedback. *Plant Cell* **14**, 611-8.

Marks, M. D. (1997). Molecular Genetic Analysis of Trichome Development in Arabidopsis. *Annu Rev Plant Physiol Plant Mol Biol* **48**, 137-163.

Masucci, J. D., Rerie, W. G., Foreman, D. R., Zhang, M., Galway, M. E., Marks, M. D. and Schiefelbein, J. W. (1996). The homeobox gene GLABRA2 is required for position-dependent cell differentiation in the root epidermis of Arabidopsis thaliana. *Development* **122**, 1253-60.

Meinhardt, H. (1994). Biological pattern formation: new observations provide support for theoretical predictions. *Bioessays* **16**, 627-32.

Meinhardt, H. and Gierer, A. (2000). Pattern formation by local self-activation and lateral inhibition. *Bioessays* **22**, 753-60.**Morohashi, K., Zhao, M., Yang, M., Read, B., Lloyd, A., Lamb, R. and Grotewold, E.** (2007) Participation of the Arabidopsis bHLH Factor GL3 in Trichome Initiation Regulatory Events *Plant Physiology* **145**, 736-746.

Nadeau, J.A., and Sack, F.D. (2002). Control of stomatal distribution on the Arabidopsis leaf surface. Science **296**, 1697–1700.

Nakamura, M., Katsumata, H., Abe, M., Yabe, N., Komeda, Y., Yamamoto, K.T., and Takahashi, T. (2006) Characterization of the Class IV Homeodomain-Leucine Zipper Gene Family in Arabidopsis. *Plant Physiology* **141**, 1363-1375.

Nakayama, M., Kikuno, R., and Ohara, O. (2002) Protein-Protein Interactions Between Large Proteins: Two-Hybrid Screening Using a Functionally Classified Library Composed of Long cDNAs. *Genome Research* **12**, 1773-1784.

Nesi, N., Debeaujon, I., Jond, C., Pelletier, G., Caboche, M., and Lepiniec, L. (2000) The TT8 gene encodes a basic helix-loop-helix domain protein required for expression of DFR and BAN genes in Arabidopsis siliques. *Plant Cell* **12**, 1863-1878.

Ohashi, Y., Oka, A., Morelli, G., and Aoyama, T. (2002) Entopically additive expression of GLABRA2 alters the frequency and spacing of trichome initiation. *Plant Journal* **29**, 359-369

Ohashi, Y., Oka, A., Rodrigues-Pousada, R., Possenti, M., Ruberti, I., Morelli, G. and Aoyama, T. (2003) Modulation of Phospholipid Signaling by GLABRA2 in Root-Hair Pattern Formation. *Science* **300**, 1427

Oppenheimer, D. G., Herman, P. L., Sivakumaran, S., Esch, J. and Marks, M. D. (1991). A myb gene required for leaf trichome differentiation in Arabidopsis is expressed in stipules. *Cell* **67**, 483-93.

Payne, C. T., Zhang, F. and Lloyd, A. M. (2000) GL3 encodes a bHLH protein that regulates trichome development in Arabidopsis through interaction with GL1 and TTG1. *Genetics* **156**, 1349-62.

Pesch, M. and Hulskamp, M. (2004). Creating a two-dimensional pattern de novo during Arabidopsis trichome and root hair initiation. *Curr Opin Genet Dev* **14**, 422-7.

Rerie, W. G., Feldmann, K. A. and Marks, M. D. (1994) The GLABRA2 gene encodes a homeo domain protein required for normal trichome development in Arabidopsis. *Genes Dev* **8**, 1388-99

Ryu, K. H., Kang, Y. H., Park, Y. H., Hwang, I., Schiefelbein, J. and Lee, M. M. (2005). The WEREWOLF MYB protein directly regulates CAPRICE transcription during cell fate specification in the Arabidopsis root epidermis. *Development* **132**, 4765-75.

Sablowski, R. W. and Meyerowitz, E. M. (1998). A homolog of NO APICAL MERISTEM is an immediate target of the floral homeotic genes APETALA3/PISTILLATA. *Cell* **92**, 93-103.

Schellmann, S., Schnittger, A., Kirik, V., Wada, T., Okada, K., Beermann, A., Thumfahrt, J., Jurgens, G. and Hulskamp, M. (2002). TRIPTYCHON and CAPRICE mediate lateral inhibition during trichome and root hair patterning in Arabidopsis. *EMBO J* **21**, 5036-46.

Schnittger, A., Jurgens, G., and Hulskamp, M. (1998) Tissue layer and organ specificity of trichome formation are regulated by GLABRA1 and TRIPTYCHON in Arabidopsis. *Development* **125**, 2283-2289.

Schnittger, A., Folkers, U., Schwab, B., Jurgens, G. and Hulskamp, M. (1999). Generation of a spacing pattern: the role of triptychon in trichome patterning in Arabidopsis. *Plant Cell* **11**, 1105-16.

Schrick, K., Nguyen, D., Karlowski, W. M. and Mayer, K. FX. (2004) START lipid/sterol-binding domains are amplified in plants and are predominantly associated with homeodomain transcription factors. *Genome Biology* **5**, R41.

Shen, B., Sinkevicius, K.W., Selinger, D.A. and Tarczynski M.C. (2006) The homeobox gene GLABRA2 affects seed oil content in Arabidopsis. *Plant Molecular Biology* **60**, 377–387.

Szymanski, D. B., Jilk, R. A., Pollock, S. M., and Marks, M. D. (1998) Control of GL2 expression in Arabidopsis leaves and trichomes. *Development* **125**, 1161-1171.

Tominaga-Wada, R., Iwata, M., Sugiyama, J., Kotake, T., Ishida, T., Yokoyama, R., Nishitani, K., Okada, K., and Wada, T. (2009) The GLABRA2 homeodomain protein directoy regulates CESA5 and XTH17 gene expression in Arabidopsis roots. *The Plant Journal* **60**, 564-574.

Uphof, J.C.T. (1962) Plant Hairs. In Encyclopedia of Plant Anatomy, Volume 5(4), W. Zimmermann and P.G. Ozenda, eds. (Berlin: Gebruder Borntrager), pp. 1-206.

Wada, T., Kurata, T., Tominaga, R., Koshino-Kimura, Y., Tachibana, T., Goto, K., Marks, M. D., Shimura, Y. and Okada, K. (2002). Role of a positive regulator of root hair development, CAPRICE, in Arabidopsis root epidermal cell differentiation. *Development* **129**, 5409-19.

Wada, T., Tachibana, T., Shimura, Y. and Okada, K. (1997). Epidermal cell differentiation in Arabidopsis determined by a Myb homolog. CPC. *Science* **277**, 1113-6.

Walker, A. R., Davison, P. A., Bolognesi-Winfield, A. C., James, C. M., Srinivasan, N., Blundell, T. L., Esch, J. J., Marks, M. D. and Gray, J. C. (1999) The TRANSPARENT TESTA GLABRA1 locus, which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein. *Plant Cell* **11**, 1337-50.

Western, T.L., Young, D.S., Dean, G.H., Tan, W.L., Samuels, A.L., and Haughn, G.W. (2004) MUCILAGE-MODIFIED4 Encodes a Putative Pectin Biosynthetic Enzyme Developmentally Regulated by APETALA2, TRANSPARENT TESTA GLABRA1, and GLABRA2 in the Arabidopsis Seed Coat. *Plant Physiology* **134**, 296-306.

Zhang, F., Gonzalez, A., Zhao, M., Payne, C. T. and Lloyd, A. (2003) A network of redundant bHLH proteins functions in all TTG1-dependent pathways of Arabidopsis. *Development* **130**, 4859-69.

Zhao, M., Morohashi, K., Hatlestad, G., Grotewold, E., and Lloyd, A. (2008) The TTG1-bHLH-MYB complex controls trichome cell fate and patterning through direct targeting of regulatory loci. *Development* **135**, 1991-1999.

Zimmermann, I. M., Heim, M. A., Weisshaar, B. and Uhrig, J. F. (2004). Comprehensive identification of Arabidopsis thaliana MYB transcription factors interacting with R/B-like BHLH proteins. *Plant J.* **40**, 22–34.

Vita

Gregory James Hatlestad was born in Minneapolis Minnesota the son of Bradley Alan Hatlestad and JoAnn Nelson Hatlestad. Following graduation from Forest Lake Senior High in Forest Lake Minnesota in 1999, he attended The University of Minnesota Twin Cities in the fall of 1999. After graduating with a Bachelor of Science degree in Biochemistry, Plant Biology and Genetics Cell Biology and Development he moved on to Institute of Cell and Molecular Biology Ph.D program at The University of Texas at Austin in 2003.

Permanent address (or email): 8057 Aenon Pl, Lino Lakes, MN, 55014

This dissertation was typed by the author.